

المملكة العربية السعودية
وزارة التعليم العالي
جامعة أم القرى
معهد خادم الحرمين الشريفين
لأبحاث الحج

التنبؤ الإحصائي لعدد الحجاج باستخدام النماذج الحديثة للسلاسل الزمنية

اعداد

د. بكري معتوق عساس　　　　د. جمال رشيد الكحلوت

ربيع الأول ١٤٢٢هـ

مقدمة

لقد حظي هذا البلد المقدس بتحمل مسؤولية تقديم الخدمات المتكاملة لحجاج بيت الله الحرام منذ أن تطأ أقدامهم أرض المملكة وحتى عودتهم الى أراضيهم سالمين غانمين . ومن هذا المنطلق ظلت الرعاية الكريمة الشاملة للقادمين لأداء فريضة الحج والعمرة موضع اهتمام أولي الأمر في المملكة العربية السعودية منذ عهد المؤسس الأول جلالة الملك عبدالعزيز رحمه الله .

ومن بيانات وزارة الحج يتضح أن أعداد الحجاج كانت ثابتة تقريباً في حدود المائة الف حاج من جميع أنحاء العالم الاسلامي في فترات السلم والازدهار الاقتصادي وكانت تنخفض الى حدود عشرين الف حاج فقط في أعوام الكسادمثل عام ١٩٢٢م والى حوالي تسعة آلاف فقط عام ١٩١٤م خلال الحرب العالمية الثانية ، ولقد بدأت الزيادة المضطردة والدائمة في أعداد الحجاج القادمين من الخارج عام ١٣٧٠هـ حيث وصل تعدادهم عام ١٣٧١هـ الى ١٤٨٥١٥ حاج أي بزيادة قدرها ٥٠٪ من العام السابق له مباشرة واستمرت هذه الزيادة حتى وصلت الى تسعة أضعاف العدد حيث وصل تعداد الحجاج القادمين من الخارج الى ٨٧٩ الف حاج عام ١٤٠١هـ والى ٩١٢ الف حاج عام ١٤٠٥هـ وظلت أعداد الحجاج تتراوح بين ٨٠٠ الى ٩٠٠ الف حاج حتى اتخاذ قرار مؤتمر وزارة خارجية الدول الاسلامية في مدينة عمان في الأردن عام ١٤٠٨هـ بأن تكون أعداد الحجاج متوافقة بنسبة واحد الى الألف لعدد السكان في الأقطار الإسلامية ومع ذلك فلا زال الحج يشهد إزدحاماً منقطع النظير لعدة أسباب منها توفير الأمن والاستقرار في مناطق الحج والانتعاش الاقتصادي لكثير من الدول الاسلامية فضلاً عن التوسع في إنشاء شبكة الطرق البرية في المملكة التي تربط المملكة بالدول المحيطة وتطوير وسائل النقل في مناطق المشاعر المقدسة إضافة الى الزيادة في أعداد السكان في البلاد الاسلامية .

ومن أهم المشاكل التي تواجه المسؤولين في المملكة العربية السعودية مشكلة نمذجة البيانات السنوية الخاصة بالحجاج بشكل علمي حديث يساهم في التنبؤ بعدد الحجاج بشكل

دقيق من الناحية الاحصائية ، ويمكن تجميع طرق النمذجة والتنبؤ بعدد الحجاج السنوي في نوعين أساسيين . النوع الأول ويعرف بنماذج التنبؤ السببية Causal Forecasting Models . ويهتم هذا النوع بتحديد المتغيرات المستقلة الأخرى المرتبطة بعدد الحجاج السنوي ثم تحديد النموذج الإحصائي الملائم الذي يصف هذه العلاقة بين المتغيرات ثم استخدام هذه العلاقة في التنبؤ . وهذا النوع من النمذجة والتنبؤ له بعض العيوب مثل ضرورة توافر بيانات تاريخية عن جميع العوامل أو المتغيرات الأخرى (وهي بالطبع كثيرة) بالإضافة الى المقدرة على التنبؤ بالعوامل المستقلة ، ويعرف هذا النوع بالتنبؤ بتحليل الانحدار Regression Analysis .أما النوع الثاني من طرق النمذجة والتنبؤ فيعرف بنماذج السلاسل الزمنية Time Series Models والذي يعتمد على تحليل البيانات التاريخية للحجاج لمعرفة وتحديد شكل ونظام البيانات Data Pattern واستكمال هذا النظام في المستقبل لإعطاء التنبؤات اللازمة بدون التعرض لأي بيانات خاصة بالمتغيرات المستقلة .

وبالرغم من أن الأساس النظري والرياضي الحديث للسلاسل الزمنية أصعب كثيراً من تحليل الانحدار ، إلا أن هذا التحليل له مميزات عديدة في دراسة ظاهرة الحج حيث يصعب حصر جميع العوامل التي تؤثر في عدد الحجاج بالإضافة الى صعوبة التنبؤ بها . ويمكن للمتخصص في مجال السلاسل الزمنية عند تحليل ظاهرة الحج تمييز طريقتين أساسيتين . الطريقة الأولى وتعرف بالطريقة التقليدية Classical Method والتي تعتمد في تطبيقها على بعض النماذج البسيطة والمحدودة جداً . وتتميز هذه الطريقة بالبساطة والسهولة الرياضية ولكن يعاب عليها أنها تفترض أن مشاهدات السلسلة الخاصة بالحج مستقلة وهذا يتعارض مع التعريف الحديث للسلاسل الزمنية مما يؤدي الى تنبؤات غير دقيقة من الناحية الاحصائية. كما أن هذه النماذج محدودة جداً . كما يعاب على هذه الطريقة أنها تفشل في كثير من الأحيان في إعطاء فترات ثقة صالحة للمشاهدات المستقبلية لعدد الحجاج . أما الطريقة الأخرى التي يمكن استخدامها في تحليل السلسلة الزمنية الخاصة بالحجاج فتعرف بطريقة بوكس وجينكنز Box-Jenkins Methodology أو ARIMA Models والتي أصبحت منذ السبعينات الأداة الأكثر قبولاً وشيوعاً في الأوساط العلمية النظرية والتطبيقية خاصة في العالم المتقدم . فقد أثبتت هذه الطريقة كفاءة عالية في النمذجة والتنبؤ بالسلاسل الزمنية .

وتتميز طريقة بوكس وجينكنز بعدة مميزات أولها أنها طريقة شاملة بمعنى أنها تعطي حلول جيدة لجميع مراحل التحليل وهي تحديد النموذج الابتدائي وتقدير معالمه وتشخيصه والتنبؤ بمشاهداته . والميزة الثانية احتوائها على مجموعة كبيرة من النماذج . والميزة الثالثة أنها لا تفترض الاستقلال بين المشاهدات وبالتالي فهي عادة أكثر كفاءة من الطريقة التقليدية . والميزة الرابعة أنها تعطي فترات ثقة ملاءمة للقيم المستقبلية . وهذه الطريقة غير منتشرة في عالمنا العربي لأن تطبيقها يحتاج الى مهارة وخبرة خاصة في مرحلة اختيار النموذج الابتدائي المناسب وتشخيصه .

ويهدف هذا البحث بصفة عامة الى استخدام اسلوب بوكس وجينكنز في إنشاء تحليل شامل للعدد السنوي للحجاج الوافدين الى المملكة ، وبصورة تفصيلية يهدف هذا البحث الى :

١- تحديد نموذج مبدئي مناسب من ضمن نماذج ARIMA لعدد الحجاج الكلي الوافد الى المملكة وذلك عن طريق تحويل السلسلة الى سلسلة مستقرة باستخدام التحويلات الرياضية ثم استخدام دالة الارتباط الذاتي Autocorrelation Function (acf) ودالة الارتباط الذاتي الجزئي Partial Autocorrelation Function (pacf) لاختيار رتب النموذج المناسب .

٢- تقدير معالم النموذج المبدئي بطريقة المربعات الصغرى غير الخطية Non linear least Squares Estimates (N.L.L.S.E) .

٣- تشخيص وفحص ملائمة النموذج المبدئي والوصول الى نموذج يحقق كل الخصائص الإحصائية المرغوب فيها .

٤- التنبؤ بالنقطة والفترة لعدد الحجاج الكلي لبعض السنوات القادمة باستخدام النموذج الذي تم تشخيصه في المرحلة السابقة .

الفصل الأول
أسلوب بوكس وجينكنز

السلسلة الزمنية هي مجموعة متتالية من المشاهدات المرتبطة المأخوذة على فترات زمنية متساوية لإحدى الظواهر مثل عدد الحجاج السنوي وعدد المواليد الشهري ودرجة الحرارة اليومية وغيرها . وتدرس السلسلة الزمنية بغرض اكتشاف نمط تاريخي معين للظاهرة ونمذجة هذا النمط واستغلال هذا النموذج في التنبؤ بالظاهرة مستقبلاً . ولقد اعتمدت الطرق التقليدية في نمذجة وتحليل السلاسل الزمنية على بعض النماذج البسيطة المحدودة والتي لا ترقى لأن تكون نظام نمذجة وتنبؤ موثوق به . فضلاً عن أن هذه الطرق التقليدية تعتمد في تطبيقها على مبدأ الاستقلال بين المشاهدات ، وهذا يتعارض مع المفهوم الأصلي للسلاسل الزمنية . كما أن هذه الأساليب تفشل في معظم الأحيان في اعطاء تنبؤات عن طريق فترات الثقة والتي تعتبر في الاحصاء أحد الأسس الهامة في الاستدلال الاحصائي .

والتحليل الحديث للسلاسل الزمنية بدأ في عام 1970 ميلادياً على يد العالمين جورج بوكس G. Box وجويلم جينكنز G. Jenkins والذي قدماه في كتابهما عام 1970 والذي يعرف في الفكر الإحصائي بإنجيل السلاسل الزمنية . ولقد أصبح تحليل السلاسل الزمنية أكثر إنتشاراً وشيوعاً في العالم خاصة المتقدم بفضل أسلوب بوكس وجينكنز الشامل والذي استطاع أن يضع حلولاً منطقية يمكن الاعتماد عليها لكل مرحلة من مراحل تحليل السلاسل الزمنية بدءاً من اختيار النموذج المبدئي وحتى مرحلة التنبؤ . ويعطي أسلوب بوكس وجينكنز نظام تنبؤ موثوق به في معظم السلاسل الزمنية التي تظهر في مجالات المعرفة المختلفة مثل البيئة والاقتصاد والفيزياء والكيمياء وغيرها .

وبالرغم من الانتشار الهائل لهذا الأسلوب الحديث والذي مازال يتصدر قائمة التحليلات الحديثة للسلاسل الزمنية ، إلا أن تطبيق هذا الأسلوب بشكل دقيق يتطلب خبرة ومهارة وممارسة قد لا تتوافر لكثير من الباحثين . وتعاني معظم الأبحاث في المنطقة العربية من بعض القصور عند تطبيق هذا الأسلوب عملياً خاصة إذا كان الباحث لا يتمتع بالخبرة والمهارة

والممارسة الكافية لاختيار النموذج المبدئي وتطوير هذا النموذج من خلال الدراسة وحتى استخدامه في التنبؤ .

ويواجه صانع القرار في المملكة العربية السعودية مشكلة وجود نظام تنبؤ موثوق به ويمكن الاعتماد عليه من خلال تنبؤات متسقة لعدد الحجاج السنوي لما تمثله هذه السلسلة من أهمية اقتصادية واجتماعية ودينية للمملكة . ولقد استخدمت العديد من الأساليب التقليدية في السلاسل الزمنية لتحليل عدد الحجاج السنوي والتي تعاني من الكثير من العيوب . ومن هنا كانت أهمية تطبيق أسلوب حديث مثل اسلوب بوكس وجينكنز لنمذجة بيانات الحج وتقدير معالم النموذج وتشخيصه واستخدامه في التنبؤ . ويعتمد اسلوب بوكس وجينكنز في تطبيقه على مجموعة أو عائلة من النماذج الحديثة تعرف باسم نماذج الإنحدار الذاتي والمتوسطات المتحركةالتكاملية ARIMA (p,d,q) والتي تساهم بشكل فعال في ايجاد تنبؤات ذات اعتمادية عالية بإستخدام تاريخ السلسلة الماضية والاحداث العشوائية التي حدثت في الماضي .

وقد يكون من المفيد قبل تحليل بيانات الحج باستخدام الاسلوب الحديث للسلاسل الزمنية أن نعطي فكرة عن المفاهيم الأساسية التي يعتمد عليها التحليل الحديث والمراحل المختلفة للتحليل بشكل نظري .

ويمكن تقسيم هذا الفصل الى :

1. المفاهيم الأساسية المستخدمة في التحليل الحديث للسلاسل الزمنية
   - الاستقرار (السكون) وأهميته .
   - دالة الارتباط الذاتي .
   - دالة الارتباط الذاتي الجزئي .
   - السلاسل الزمنية غير المستقرة .
   - بعض التحويلات الرياضية لتسكين السلسلة .

2. نماذج السلاسل الزمنية الحديثة
- نماذج الانحدار الذاتي AR(P)
- نماذج المتوسطات المتحركة MA(q)
- نماذج ARMA المختلطة .

3. طريقة بوكس وجينكنز في التحليل الحديث
- تحديد النموذج المبدئي .
- تقدير المعالم وتشخيص النموذج .
- التنبؤ بالمشاهدات المستقبلية .

1. الاستقرار (السكون) Stationarity

يقال أن السلسلة الزمنية y_t مستقرة Stationary إذا كانت الخصائص الأحصائية الأساسية (التوقع والتباين والتغاير) للسلسلة ثابتة خلال الزمن . أي أن السلسلة الزمنية y_t تكون مستقرة في التوقع والتباين والتغاير إذا توافرت جميع الشروط الآتية :

(i) $E(y_1) = E(y_2) = \dots = E(y_n) = \mu$

(ii) $V(y_1) = V(y_2) = \dots = V(y_n) = \sigma^2$

(iii) $Cov(y_t, y_{t+k}) = \gamma(\kappa)$

حيث μ , σ^2 قيم ثابتة لا تعتمد على t . والدالة $\gamma(k)$ تعتمد على البعد الزمني بين y_t , y_{t+k} فقط ولا تعتمد على الزمن t . أي أن

$Cov(y_1, y_2) = Cov(y_2, y_3) = \dots Cov(y_{t,t+1}) = \gamma(1)$

$Cov(y_1, y_3) = Cov(y_2, y_4) = \dots Cov(y_{t,t+2}) = \gamma(2)$

وهكـــــذا .

ويتضح من تعريف الاستقرار أن معظم الظواهر الاقتصادية والطبيعية والبيئية وغيرهـــــا

تعاني من عدم الاستقرار في أحد خصائصها الأساسية على الأقل . وسنرى فيما بعد أن تطبيق الأسلوب الحديث للسلاسل الزمنية يتطلب تحويل السلسلة غير المستقرة الى سلسلة مستقرة قبل التحليل باستخدام بعض التحويلات الرياضية أو ما يعرف بتسكين السلسلة .

2. دالة الارتباط الذاتي Autocorrelation Function

إذا كان لدينا السلسلة الزمنية الآتية $y_1, y_2, \dots, y_n$ فإننا نعرف $\rho(K)$ بأنها الارتباط الذاتي بين مشاهدات السلسلة التي تبعد عن بعضها البعض مسافة مقدارها k وحدات زمنية . فمثلاً $\rho(1)$ تعني الارتباط بين المشاهدات التي تبعد عن بعضها البعض بمقدار وحدة زمنية واحدة أي الارتباط بين (y_1, y_2) مثلاً أو بين (y_{10}, y_{11}) ، ... أو الارتباط بين (y_t, y_{t+1}) بصفة عامة . و $\rho(3)$ تمثل الارتباط بين المشاهدات التي تبعد عن بعضها البعض بمقدار ثلاث وحدات زمنية أي الارتباط بين (y_1, y_4) أو بين (y_2, y_5) ، ... أو الارتباط بين (y_t, y_{t+3}) بصفة عامة ... وهكذا .

فإذا كانت السلسلة مستقرة فإن

$$\rho(1) = Cov(y_t, y_{t+1}) / V(y_t) = \gamma(1) / \gamma(0)$$
$$\rho(2) = Cov(y_t, y_{t+2}) / V(y_t) = \gamma(2) / \gamma(0)$$

وبصفة عامة

$$\rho(k) = Cov(y_t, y_{t+k}) / V(y_t) = \gamma(k) / \gamma(0)$$

وبذلك يمكن تمثيل الدالة $\rho(\kappa)$ ضد قيم K بيانياً. وسنرى فيما بعد أن دالة الارتباط الذاتي من أهم الأدوات التي تلعب دوراً هاماً في التحليل الحديث للسلاسل الزمنية .

3. دالة الارتباط الذاتي الجزئي Partial Autocorrelation Function

في موضوعات السلاسل الزمنية قد نجد أن جزء من العلاقة بين المتغيرين (y_1, y_3) يعود الى ارتباط كل من هذين المتغيرين بالمتغير الأخر y_2 الذي يقع بينهما . وقد يكون من المفيد

أحياناً أن نبحث في العلاقة بين (y_1 , y_3) بشرط ثبات المتغير الذي بينهما وهو y_2 أي نحاول البحث في العلاقة بين (y_1 , y_3) بعد حذف تأثير y_2 على المتغيرين . ويعرف هذا الارتباط بالارتباط الجزئي من الرتبة الثانية ويرمز له عادة بالرمز ϕ_{22} . وبصفة عامة يمكن تعريف دالة الارتباط الجزئي ϕ_{kk} بأنها الدالة التي تعطي العلاقة بين (y_t , y_{t+k}) بعد حذف تأثير المتغيرات التي بينهما وهي $y_{t+1} , y_{t+2}, \ldots, y_{t+k-1}$. وسنرى أيضاً أن هذه الدالة تلعب دوراً أساسياً في بناء نماذج السلاسل الزمنية الحديثة .

4. السلاسل الزمنية غير المستقرة Non Stationary Time Series

تتصف معظم السلاسل الزمنية التي تصادفنا في الحياة العملية بخاصية عدم الاستقرار (عدم السكون) ، وهذه الخاصية تجعل تحليل السلاسل الزمنية والتنبؤ بها أمر في غاية الصعوبة . ولحسن الحظ فإن معظم الظواهر أو السلاسل الزمنية التي تنشأ في الاقتصاد والادارة والهندسة والطبيعة والكيمياء وعلوم البيئة تتصف بخاصية التجانس . ويقصد بهذه الخاصية أنه بالرغم من أن السلسلة الأصلية y_t قد تكون غير مستقرة فإنه يمكن تحويل هذه السلسلة الى سلسلة أخرى مستقرة Z_t باستخدام بعض التحويلات الرياضية . وهذا النوع من السلاسل الذي يمكن تحويله الى سلاسل مستقرة يطلق عليه السلاسل الزمنية المتجانسة Homogenous Time Series .

5. بعض التحويلات الرياضية لتسكين السلاسل الزمنية

يمكن تحويل معظم السلاسل الزمنية غير المستقرة الى سلاسل مستقرة عن طريق بعض التحويلات الرياضية البسيطة . وفيما يلي نستعرض أهم هذه التحويلات بشكل سريع :

(a) أخذ فروق السلسلة

والمقصود بأخذ الفروق هو تحويل السلسلة الزمنية الأصلية y_t الى سلسلة جديدة z_t عن طريق التحويل الآتي :

$$z_t = \Delta y_t = y_t - y_{t-1} , \qquad t = 2,3,\ldots, n \qquad (1)$$

وتسمى السلسلة الجديدة z_t بسلسلة الفروق الأولى للسلسلة y_t وعدد مشاهدات هذه السلسلة الجديدة هو (n-1) أي أننا نفقد مشاهدة واحدة عند أخذ الفروق الأولى في سبيل الوصول الى سلسلة مستقرة . وبالرغم من بساطة هذا التحويل ، إلا أنه مفيد جداً في تحويل كثير من السلاسل الى سلاسل مستقرة .

وقد تظل سلسلة الفروق الأولى z_t غير مستقرة وهنا يجب أن نأخذ الفروق الثانية أي $\Delta^2 y_t$ أو Δz_t وهذا النوع مفيد جداً في تحويل السلاسل الى سلاسل مستقرة ويمكن تصور الفروق الثانية w_t كالأتي :

$$w_t = \Delta^2 y_t = \Delta z_t = z_t - z_{t-1} \; , \qquad t = 3,4,\dots, n$$

وعدد مشاهدات السلسلة الجديدة w_t هو (n-2) أي أننا نفقد مشاهدتين عند أخذ الفروق الثانية للسلسلة y_t .

(b) أخذ فروق اللوغاريتمات

وقد تكون سلسلة الفروق الثانية w_t غير مستقرة أيضاً ، وهنا لابد للباحث أن يحاول نوع آخر من التحويلات خاصة إذا كان تباين السلسلة الأصلية غير مستقر . ويعرف هذا النوع من التحويلات بفروق اللوغاريتمات . أي نحول السلسلة الأصلية y_t الى سلسلة أخرى جديدة y_t^* كالآتي :

$$y_t^* = \log y_t \qquad , t = 1,2,3,\dots,n \qquad (2)$$

ثم أخذ فروق اللوغاريتمات كالآتي

$$z_t^* = y_t^* - y_{t-1}^* = \log y_t - \log y_{t-1} \qquad \dots (3)$$
$$; \; t = 2,3, \dots , n$$

2. نماذج السلاسل الزمنية الحديثة

تعتمد النماذج الحديثة للسلاسل الزمنية على المبدأ الفلسفي القائل بأن الحاضر هو نتاج

الماضي ، وهذا يعني التعبير عن قيمة السلسلة عند الزمن t أي y_t كدالة في القيم الماضية للسلسلة أي y_{t-1} , y_{t-2} , ... بالأضافة الى التغيرات العشوائية التي حدثت في الماضي أي ε_{t-1} , ε_{t-2} , ... ويضاف الى هذا النموذج مركبة الخطأ العشوائي عند الزمن t أي ε_t.
أي أن النماذج الحديثة للسلاسل الزمنية تفترض أن

$$y_t = f(y_{t-1}, y_{t-2}, \dots, \varepsilon_{t-1}, \varepsilon_{t-2} \dots) + \varepsilon_t$$

ومن أشهر هذه النماذج مايعرف باسم نماذج ARMA (p, q) الخطية والتي كانت أساساً لطريقة بوكس وجينكنز ، ويمكن كتابة هذه النماذج بصورة عامة على الشكل

$$y_t = \sum_{i=1}^{p} \phi_i \, y_{t-i} - \sum_{j=1}^{q} \theta_j \, \varepsilon_{t-j} + \varepsilon_t \qquad \dots (1)$$

ويمكن تمييز ثلاثة نماذج أساسية من هذه العائلة وهي :

- نماذج الانحدار الذاتي أو AR(p)
- نماذج المتوسطات المتحركة أو MA(q)
- النماذج المختلطة أو ARMA (p, q)

1. نماذج الانحدار الذاتي Autoregressive Models AR (p)

ويقصد به التعبير عن المشاهدة الحالية للسلسلة الزمنية y_t كدالة خطية في المشاهدات السابقة بالأضافة الى متغير عشوائي يرمز له بالرمز ε_t أي

$$y_t = \phi_1 \, y_{t-1} + \phi_2 \, y_{t-2} + \dots + \phi_p \, y_{t-p} + \varepsilon_t \qquad \dots (2)$$

حيث ε_t متغيرات عشوائية مستقلة تتبع توزيع طبيعي توقعه الصفر وتباينه σ^2 ثابت وأنها مستقلة عن y_{t-1} , y_{t-2} , ... وأحياناً ما يطلق على المتغيرات العشوائية ε_t بأنها White noise .

وهناك بعض القيود التي توضع على المعالم لضمان استقرار هذا النوع من النماذج .

فمثلاً النموذج AR (1) يكون مستقر إذا كان $|\phi_1| < 1$. وفي هذه الحالة نجد أن دالة الارتباط الذاتي لهذا النموذج تأخذ الشكل الآتي

$$\rho(k) = \phi_1^k \qquad ; k = o, 1, 2, \ldots \qquad (3)$$

وهي من الملامح الأساسية لهذا النموذج وتعرف بأنها دالة الذاكرة .

ويمكن التعبير عن نموذج AR (1) كالآتي

$$y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \ldots$$

أي يمكن التعبير عن y_t بدلالة التغير العشوائي الحالي بالأضافة الى عدد لا نهائي من الحدود من التغيرات العشوائية السابقة . ولذلك يقال أن هذا النموذج له ذاكرة لا نهائية ولكن هذه الذاكرة تختفي تدريجياً .

وتتميز دالة الارتباط الذاتي لنماذج AR (p) المستقرة بالتناقص السريع بشكل أسي أو في صورة موجات من دالة الجيب Sine waves . أما دالة الارتباط الذاتي الجزئي للنموذج AR (p) فتنقطع كلية بعد المسافة p .

2. نماذج المتوسطات المتحركة Moving Average Models MA (q)

ويمكن التعبير عن هذا النموذج في الشكل التالي :

$$y_t = -\theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t \qquad (4)$$

حيث ε_t لها نفس الخصائص التي سبق أن ذكرناها في حالة نماذج AR (p) . وتنقطع دالة الارتباط الذاتي لهذا النوع من النماذج بعد الرتبة q. أما دالة الارتباط الذاتي الجزئي فهي تتناقص بشكل أسي أو موجات من دالة الجيب . ولذلك تأخذ سلوك هاتين الدالتين سلوك معاكس لحالة نماذج AR (p) .

3. النماذج المختلطة ARMA (p , q) Models

ويمكن التعبير عن هذه النماذج في الصورة

$$y_t = \sum_{i=1}^{p} \theta_i \, y_{t-i} - \sum_{i=1}^{q} \theta_i \, \varepsilon_{t-i} + \varepsilon_t \qquad \ldots (5)$$

ودالة الارتباط الذاتي $\rho(k)$ لهذا النوع من النماذج تتناقص بشكل أسي أو موجات من دالة الجيب . ويتشابه سلوك دالة الارتباط الذاتي الجزئي ϕ_{kk} مع سلوك دالة الارتباط الذاتي الى حد كبير .

3. طريقة بوكس وجينكنز

تعرف طريقة بوكس وجينكنز في تحليل السلاسل الزمنية بالطريقة الحديثة في تحليل السلاسل الزمنية . وقد اكتملت أركان هذه الطريقة من نظريات احصائية ووسائل حسابية في السبعينات من القرن العشرين . وتعتبر هذه الطريقة نقله نوعية متميزة في تحليل السلاسل الزمنية بحيث أصبحت أكثر الطرق شيوعاً وتفضيلاً من قبل الدارسين والباحثين لأسباب عديدة. والسبب الأول في هذا الشيوع والتفضيل أنها طريقة شاملة بمعنى أنها تقدم حلول مناسبة لجميع مراحل تحليل السلاسل الزمنية . والسبب الثاني أن هذه الطريقة تعتمد على نماذج ARMA والتي أظهرت نجاحات متعددة في النمذجة والتنبؤ . والسبب الثالث أن هذه الطريقة تعتبر نظام نمذجة وتنبؤ موثوق به ويمكن الاعتماد عليه في اعطاء تنبؤات متسقة إحصائياً . والسبب الرابع أنها تمد الباحث بفترات ثقة مناسبة لمعالم النموذج وللمشاهدات المستقبلية . والسبب الأخير هو توافر العديد من الحزم الاحصائية القادرة على تنفيذ جميع مراحل التحليل بكفاءة عالية مثل SAS , SPSS , MINITAB , TSERIES وغيرها .

ولا يعني هذا أن طريقة بوكس وجينكنز هي الطريقة المثالية في تحليل السلاسل الزمنية . فمثل هذه الطريقة (المثالية) غير موجود في مجال السلاسل الزمنية حتى الآن وربما كان هذا هو السبب في تجديد خلايا هذا المجال كل يوم بطرق أكثر حداثة في أحد موضوعات تحليل السلاسل الزمنية . وربما العيب الأساسي في هذه الطريقة أنها تحتاج الى مهارات وخبرات

خاصة من قبل الباحث لتطبيقها عملياً . وإذا وجد مثل هذا الباحث فإن الطريقة تعطي نتائج ممتازة .

ويتكون أسلوب بوكس وجينكنز من أربع مراحل أساسية هي :

١- تحديد النموذج المبدئي للبيانات (تعريف النموذج)

٢- تقدير معالم النموذج المبدئي .

٣- تشخيص كفاءة النموذج

٤- التنبؤ بالمشاهدات المستقبلية

وفيما يلي نقدم عرضاً مبسطاً لكل مرحلة من هذه المراحل الأربع .

1. تحديد النموذج المبدئي (التعريف) Identification

ويتم في هذه المرحلة التعرف على النموذج المبدئي الذي يلائم البيانات . ويقصد بالتعرف على النموذج هو تحديد أو اختيار رتب النموذج p و q المناسبة للبيانات المتاحة ، أي اختيار نموذج ابتدائي من بين عائلة نماذج ARMA (p, q) . وتعتبر هذه المرحلة أهم وأصعب مراحل التحليل ليس في مجال السلاسل الزمنية فقط ولكن في جميع مجالات الاحصاء . وتتطلب هذه المرحلة مهارة وخبرة وممارسة عملية من الباحث بالأضافة الى حكمه الشخصي . وهناك خطوط عريضة لهذه المرحلة يمكن تلخيصها فيما يلي :

أولاً : التأكد من استقرار السلسلة

ويتم ذلك عادة بالتوقيع البياني للسلسلة بالأضافة الى ايجاد ورسم دالة الارتباط الذاتي للسلسلة $r(k) = \hat{\rho}(k)$. فإذا تناقصت هذه الدالة بسرعة دل ذلك على استقرار السلسلة y_t أما إذا تناقصت بشكل بطئ فهذا دليل على عدم استقرار السلسلة ويجب تحويلها الى سلسلة أخرى مستقرة بإستخدام بعض التحويلات الشهيرة مثل الفروق أو فروق اللوغاريتمات أو الجذر التربيعي ... الخ . بمعنى آخر اننا هنا يجب أن نحدد أصغر قيمة للفروق اللازمة لتسكين السلسلة وترمز لهذه القيمة بالرمز d .

ثانياً : ايجاد دوال الارتباط الذاتي $r(k)$ والذاتي الجزئي $\hat{\phi}_{kk}$ للسلسلة المستقرة .

ثالثاً : مقارنة رسم الدوال $r(k)$, $\hat{\phi}_{kk}$ للبيانات بالرسوم النظرية للدوال $\rho(k)$, ϕ_{kk} والتي تميز الملامح الأساسية لنماذج ARMA المختلفة .

رابعاً : من المقارنة السابقة يمكن التعرف على النموذج الملائم للبيانات أي يمكن معرفة قيم p و q الابتدائية المناسبة للبيانات .

ونحن نعلم من دراستنا لنظرية الإحصاء أن هناك بعض الفروق بين الخصائص النظرية الرياضية (خصائص المجتمع) (المعالم) وخصائص العينة المناظرة (الاحصاءات) (التقديرات). وبالتالي فمن المتوقع أن نجد إختلافاً بين الرسوم النظرية للدوال $\rho(k)$, ϕ_{kk} والتي تميز العمليات العشوائية الأساسية للبيانات والدوال $r(k)$, $\hat{\phi}_{kk}$ المحسوبة للعينة أو للسلسلة . فعلى سبيل المثال من المعروف نظرياً أن دالة الارتباط الذاتي النظرية $\rho(k)$ لنموذج المتوسطات المتحركة ذات الرتبة الأولى (MA (1)) تنقطع كلية بعد المسافة الأولى من الناحية النظرية . ولكن إذا كان لدينا سلسلة أو عينة حقيقية ذات حجم معين n تنتمي بالفعل لهذا النموذج فإننا نتوقع أن يكون رسم الدالة $r(k)$ لهذه السلسلة قريباً من الشكل النظري أو قريباً من نظامه ولا ينطبق عليه تماماً .

وبالتالي لابد للباحث من توظيف أدوات نظرية الاحصاء خاصة فيما يتعلق بموضوع إختبارات الفروض لدراسة معنوية دوال الارتباط الذاتي $\rho(k)$ والارتباط الذاتي الجزئي ϕ_{kk} وذلك لمعرفة النموذج الابتدائي الملائم للبيانات المتاحة . وبالتالي لابد للباحث أن يكون قادراً على إختبار الفروض الآتية :

(i) $H_o : \rho(k) = o ; \quad H_1 : \rho(k) \neq o , k = 1,2,\ldots$

(ii) $H_o : \phi_{kk} = o ; \quad H_1 : \phi_{kk} \neq o , k = 1,2,\ldots$

ولإجراء مثل هذه الاختبارات لابد أن يكون لدينا إحصاءات مناسبة معروف توزيعها الاحتمالي . ولقد أثبت بارتليت Barttlet أن الأخطاء المعيارية للمقدرات r (k) ممكن تقريبها بالصور الآتية :

$$SE(r(1)) \approx \sqrt{\frac{1}{n}}$$

$$SE(r(k)) \approx \sqrt{\frac{1}{n}[1+2\sum_{i=1}^{k-1} r^2(i)]} \qquad , \; k \geq 2$$

ولقد أثبت أندرسون Anderson أن الاحصاء $Z = r(k) / S.E(r(k))$ يتبع تقريباً توزيع طبيعي قياسي إذا كان فرض العدم صحيح . وبالتالي نرفض H_0 إذا كان $|Z| \geq z_{\alpha/2}$.

وعادة مايستخدم الباحث مستوى المعنوية $\alpha = 0.05$ وبالتالي يستدل على أن $\rho(k)$ لا يختلف معنوياً عن الصفر إذا كان

$$-2\,SE(r(k)) \leq r(k) \leq 2\,SE(r(k))$$

إلا أن الخبراء في مجال السلاسل الزمنية يحذرون من تثبيت قيمة α لكل الاختبارات وبالتالي قد يكون من المفيد تغيير قيمة α من إختبار لآخر . وعلى العموم فإن اختيار قيم α المناسبة يحكمها خبرة ومهارة الباحث وحكمه الشخصي والذي يستطيع أن يحدد نظام دوال الارتباط الذاتي والذاتي الجزئي بفحص الدوال المناظرة للبيانات المتاحة . وعادة ماتستخدم قيم كبيرة لمستوى المعنوية α عند إجراء الاختبارات للمسافات الصغيرة مثل $\rho(1)$ و $\rho(2)$ وتستخدم مستوى معنوية أصغر للمسافات البعيدة . وهنا نحذر مرة ثانية من استخدام نتائج الحاسب الآلي فيما يتعلق بهذه الاختبارات والتي غالباً ماتستخدم مستوى معنوية 5% لاجراء مثل هذه الاختبارات .

أما بالنسبة لدالة الارتباط الذاتي الجزئي فإننا ممكن أن نستخدم الاحصاء الأتي :

$$Z = \hat{\phi}_{kk}\sqrt{n}$$

والذي يتبع تقريباً توزيع طبيعي قياسي بإفتراض صحة الفرض العدمي $H_0 : \phi_{kk} = 0$.

وبالتالي نرفض H_0 إذا كان

$$|Z| \geq z_{\alpha/2}$$

حيث تعرف القيمة $z_{\alpha/2}$

$$P [Z \geq z_{\alpha/2}] = \alpha/2$$

مرة أخرى نحذر من تثبيت قيمة α لكل الاختبارات .

وبناء على نتائج الاختبارات المختلفة من هذه الأنواع يمكن للباحث تحديد الرتب المناسبة q, p . وبالتالي يكون الباحث قد تعرف على نموذج ARIMA (p, d, q) المبدئي الملائم للبيانات .

2. تقدير معالم النموذج

بعد تحديد النموذج المبدئي المناسب باختيار قيم p, d, q تأتي مرحلة تقدير معالم هذا النموذج . وتستخدم في هذه المرحلة طريقة المربعات الصغرى غير الخطية للمعالم . ولابد من استخدام الطرق العددية لايجاد مثل هذه التقديرات . وكل الحزم الاحصائية التي تحتوي على التحليل الحديث للسلاسل الزمنية تمد الباحث بمثل هذه التقديرات .

3. تشخيص النموذج

بمجرد إختيار النموذج الابتدائي وتقدير معالمه يجب دراسة كفاءة هذا النموذج وملائمة فروضه المختلفة ودراسة امكانية تحسينه وتطويره . وهذه المرحلة من أهم وأصعب المراحل وترتبط ارتباط وثيق بجودة إختيار النموذج المبدئي . ونحذر هنا من استخدام R^2 أو مربع معامل التحديد . فالدور الذي يلعبه R^2 في مجال السلاسل الزمنية محدود جداً ولا نستطيع استخدامه للاختيار بين النماذج المختلفة . وتقوم مرحلة تشخيص النموذج على ثلاث مجموعات لفحص مدى ملائمة النموذج المبدئي وهي :

A. تحليل البواقي Residual Analysis

إذا كان النموذج المبدئي يمثل بالفعل حقيقة عملية ARIMA التي ولدت بيانات السلسلة المتاحة فإن البواقي e_t الناتجة يجب أن تحقق الفروض الموضوعة والخاصة بالتغيرات العشوائية ε_t وهي أن تتبع توزيع طبيعي متوسطه الصفر وتباينها ثابت وغير مرتبطة مع بعضها زمنياً، كما أنها تكون غير مرتبطة مع المتغيرات المستقلة $\ldots \varepsilon_{t-1}, \varepsilon_{t-2}$ $y_{t-1}, y_{t-2}, \ldots,$. ومن ثم يجب التحقق على الأقل من أن هذه البواقي لا تظهر خلل واضح في أي من هذه الفروض . وفي الواقع أن هناك العديد من الوسائل لتحقيق هذه الأهداف نذكر منها الأتي :

1. رسم البواقي e_t كسلسلة زمنية أي ضد الزمن t . وهذا الرسم قد يظهر بعض مصادر عدم الاستقرار مثل عدم ثبات التباين . كما يساعد هذا الرسم بشكل فعال على اكتشاف أي قيم شاذة التي قد توجد لأي سبب .

2. رسم دالة الارتباط الذاتي للبواقي للتأكد من أن ليس هناك أي قيم معنوية . وبالطبع يمكن استخدام صيغة بارتلت لاكتشاف القيم المعنوية . ويلاحظ أن هذا التشخيص خاص بكل قيمة من معاملات الارتباط الذاتي ويجري بشكل مستقل من معامل الى أخر .

3. إختبار Box - Pierce

إذا كان النموذج المبدئي صحيح فإن الأخطاء يجب أن تكون عشوائية بحتة ولها التوزيع الطبيعي بمتوسط صفر وتباين ثابت . وبالتالي فإن مقدرات الارتباطات الذاتية r(k) للبواقي تكون غير معنوية وقريبة في توزيعها من التوزيع الطبيعي بمتوسط صفر وتباين ثابت يساوي n^{-1} تقريباً . وبالتالي فإن المقدر

$$Q = n \sum_{i=1}^{k} r_e^2 \ (i)$$

يتبع توزيع χ^2 بدرجات حرية (k - p -q) إذا كان النموذج وفروضه مناسبة . وبالتالي إذا كانت قيمة Q أقل من القيمة الجدولية χ^2_α نقبل ملائمة النموذج حيث تعرف χ^2_α كالتالي :

$$P\,[\chi^2_{(k-p-q)} > \chi^2_\alpha] = \alpha$$

حيث α مستوى المعنوية .

4. تقويم معاملات الارتباط الذاتي للفروق الأولى للبواقي

إذا كانت الأخطاء ε_t عشوائية بحتة فإن الفروق الأولى لها يجب أن تتبع توزيع MA(1) بمعلمة $\theta_1=1$. كما أن معامل الارتباط الذاتي من الرتبة الأولى للبواقي يجب أن يساوي 0.5- فإذا كانت نتائج تحليل البواقي لا تظهر أي خلل حقيقي لهذه الملامح فقد يكون هذا دليل أخر على أن الأخطاء تتبع تغيرات عشوائية بحتة .

5. رسم البواقي e_t ضد المتغيرات المستقلة $y_{t-1}, y_{t-2}, \ldots, \varepsilon_{t-1}, \ldots$ يجب ألا يظهر أي ارتباط واضح .

6. تحليل البواقي e_t يجب ألا يظهر أي خلل في فرض اعتيادية الاخطاء ε_t .

إذا كان الاختبار Q أو أحد الاختبارات الأخرى التي تستخدم في تشخيص النموذج المبدئي أدى الى عدم ملائمة هذا النموذج فإن الباحث يجب أن يعود الى المرحلة الأولى لاختيار نموذج أخر ثم يقوم بتقدير معالمه وتشخيصه أيضاً قبل الانتقال الى مرحلة التنبؤ . وقد تستمر عملية تشخيص كفاءة النموذج عدة مرات حتى نصل الى نموذج كفء يستخدم في التنبؤ .

B. إضافة بعض المعالم Overfitting

إذا كانت القيم المختارة للمعالم p و q في المرحلة الأولى هي p_0 و q_0 فإنه يمكن توفيق النموذج الأعلى مباشرة في الرتبة في اتجاهين . مرة بإضافة معلمة من معالم الانحدار الذاتي وبالتالي يمكن توفيق النموذج $ARMA\ (p_0 + 1, q_0)$ ، ومرة بإضافة معلمة من معالم المتوسطات المتحركة أي بتوفيق النموذج $ARMA\ (p_0 , q_0 + 1)$. ثم إختبار معنوية المعلمة المضافة ، فإذا كانت المعلمة المضافة في كلا الاتجاهين لا تختلف معنوياً عن الصفر فقد يكون هذا دليل على كفاءة وملائمة النموذج المبدئي $ARMA\ (p_0 , q_0)$. أما إذا كانت إحدى هاتين المعلمتين معنوية ، فقد يكون هذا دليل على أن النموذج الأعلى أكثر ملائمة من

النموذج المبدئي .

C. حذف بعض المعالم من النموذج Underfitting

بعد تقدير معالم النموذج المبدئي يجب على الباحث حساب تقديرات الأخطاء المعيارية لمقدرات المعالم وحذف أي معالم غير ضرورية من النموذج ثم توفيق النموذج الأدنى إن وجدت مثل هذه المعالم . ويتم إختبار المعنوية عن طريق قيم الاحصاء T . ووجود مقدر لا يختلف معنوياً عن الصفر قد يعتبر مؤشراً جيداً على إمكانية تبسيط هذا النموذج بتخفيض عدد معالمه عن طريق توفيق النموذج ذات الرتبة الأدنى إذا كانت رتبة هذا المقدر هي أعلى رتبة في النموذج .

4. التنبـــــؤ Forecasting

التنبؤ هو المرحلة الأخيرة في تحليل السلاسل الزمنية ، وهو عادة الهدف النهائي من تحليل السلاسل الزمنية . ولكي يكون نظام التنبؤ موثوق به لابد أن يكون الباحث على قناعة تامة بأن النموذج الذي تم إختياره وتشخيصه وتقدير معالمه هو النموذج الأكثر كفاءة وملائمة للبيانات المتاحة . والتنبؤ الجيد هو الذي يعتمد على نموذج ذو كفاءة عالية أي نموذج استطاع أن يمر بنجاح عبر مراحل التشخيص السابق ذكرها . إذا كانت السلسلة الزمنية $y_1, y_2, \dots, y_n$ ونريد التنبؤ بالقيمة المستقبلية y_{n+h} أي القيمة التي ستحدث بعد h من الفترات الزمنية ، فإن اسلوب بوكس وجينكنز يعتمد على تقدير القيمة المتوقعة للمتغير العشوائي Y_{n+h} باستخدام المعلومات المتاحة حتى الزمن n . ونلاحظ في تقدير هذه القيمة المتوقعة مايلي :

1. يتم استبدال الاخطاء الحالية ε_t بالبواقي الفعلية e_t .
2. يتم استبدال الاخطاء المستقبلية $\varepsilon_{t+j}\,;\ o < j \le h$ أي الأخطاء التي لم تحدث بعد بقيمها المتوقعة أي الصفر .
3. يتم استبدال المشاهدات اللاحقة y_{n+j} $(o < j < h)$ بقيمها التنبؤية $\hat{y}_{n+j}$.

ولكي يتم التنبؤ بالقيم y_{n+h} لابد أن يسبقه التنبؤ بالقيم y_{n+h-1} , ... y_{n+2}، y_{n+1} أولاً . ويعتبر نظام بوكس وجينكنز في التنبؤ مثالي لأن التنبؤات تناظر أقل متوسط مربعات خطأ ممكن . وهذه التنبؤات من أهم مميزات أسلوب بوكس وجينكنز .

ومن المميزات الأخرى لأسلوب بوكس وجينكنز أنه يعطي فترات ثقة عالية الاعتمادية للمشاهدات المستقبلية سواء كانت النماذج موسمية أو غير موسمية . ولكي يتم بناء فترات ثقة لهذه المشاهدات يتم عادة كتابة النموذج الذي تم تشخيصه في صورة متوسطات متحركة كالأتي :

$$y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots$$

وبالتالي يمكن تقريب الأخطاء المعيارية للمقدرات المستقبلية كما يلي :

$$SE(\hat{y}_{n+1}) \approx S = \sqrt{MSE}$$

$$SE(\hat{y}_{n+h}) \approx S\sqrt{1 + \sum_{i=1}^{h-1} \psi_i^2} \quad , \quad h > 1$$

وباستخدام التوزيع الطبيعي كتقريب للمقدرات المستقبلية فإن $100(1 - \alpha)\ \%$ فترة ثقة للمشاهدة y_{n+h} هي :

$$\hat{y}_{n+h} \pm z_{\alpha/2}\, SE(\hat{y}_{n+h})$$

الفصل الثاني
تحليل بيانات الحج

1. الفحص الأولي للبيانات

تتكون السلسلة الزمنية لعدد الحجاج الوافدين الى المملكة العربية السعودية تحت الدراسة من 25 مشاهدة تمثل تطور الحجاج من سنة ١٣٩٦هـ الى سنة ١٤٢٠هـ. والخطوة الأولى الهامة في تحليل بيانات الحج هو رسم السلسلة الزمنية بيانياً وذلك لأن هذا الرسم البياني يظهر الملامح الأساسية للبيانات موضع الدراسة مثل الاتجاه العام والبيانات الشاذة والتشتت. ولقد تم بالفعل توقيع السلسلة بيانياً كما هو واضح في الشكل (1) . وبفحص الرسم البياني نلاحظ وجود إتجاه عام للسلسلة للزيادة خلال الفترة الزمنية للدراسة ، وبذلك تكون السلسلة غير مستقرة في متوسطها الحسابي . وتعود الزيادة في متوسط عدد الحجاج الى عدة أسباب مثل:

1. تطور حجم المجتمعات الاسلامية .
2. تطور مستوى معيشة المجتمعات الاسلامية .
3. تطور وسائل النقل الحديثة .
4. الخدمات والتسهيلات التي تقدمها المملكة الى الحجاج .
5. التوسعات في الحرمين وتطوير شبكة الطرق التي تمت في هذه الفترة الزمنية .

كما يلاحظ من الفحص البياني لعدد الحجاج تغير التشتت في بيانات السلسلة مع مرور الزمن ، وهذا يعني أن تباين المشاهدات أيضاً غير مستقر مع الزمن . ويمكن من خلال الفحص البياني إكتشاف بعض القيم غير العادية (الشاذة) مثل عدد الحجاج في عام ١٤٠٣هـ أي عند الوحدة الزمنية الثامنة والذي تميز بوثبة فجائية الى أعلى . أيضاً الفترة الزمنية من عام ١٤٠٨هـ الى عام ١٤١١هـ (أي عند الوحدات الزمنية $t = 13, 14, 15,16$) تميزت بإنخفاض ملحوظ في عدد الحجاج ووصل الى أدنى مستوى له عند عام ١٤١١هـ (أي عند الوحدة الزمنية السادسة عشرة) . ولذلك كان من الضروري العودة الى مصدر البيانات الأصلية للتأكد من صحة هذه البيانات ، وبالفعل وجد أن هذه البيانات دقيقة وليس بها أي أخطاء .

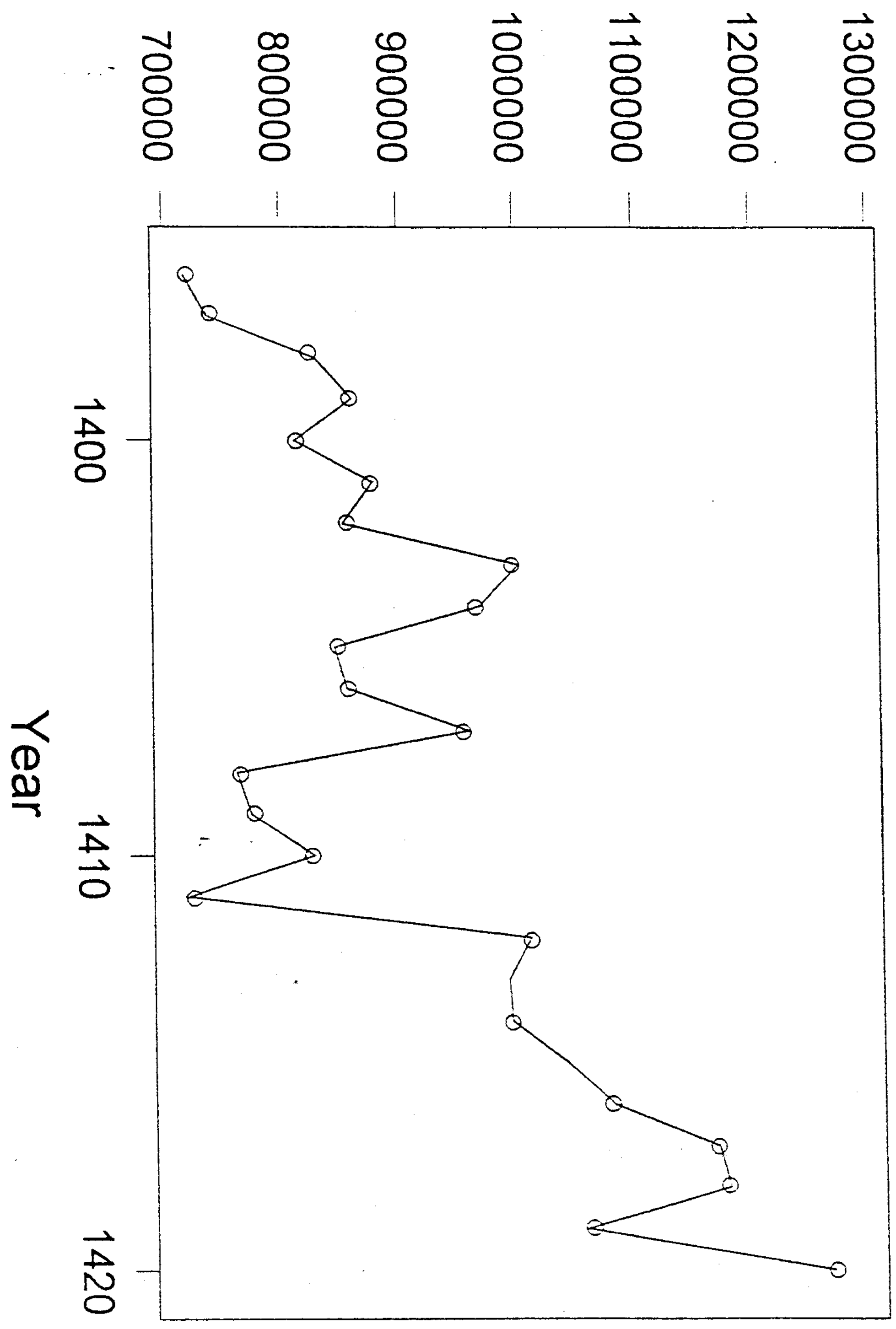

Time Series Plot Of Pilgrims From Year 1396 To 1420
Pilgrims
700000
800000
900000
1000000
1100000
1200000
1300000
1400
1410
1420
Year

Fig.(1)

وبإختصار شديد أظهر الوصف والتحليل المبدئي لسلسلة الحجاج الخصائص الأتية :

1. عدم استقرار المتوسط الحسابي .
2. عدم إستقرار التباين .
3. وجود بعض القيم غير العادية .

ونتيجة لعدم استقرار الخصائص الإحصائية الهامة للسلسلة كان من الضروري تحويل هذه البيانات الأصلية الى بيانات أو سلسلة أخرى مستقرة . ولكن قبل تحويل البيانات والحكم النهائي بعدم استقرار السلسلة y_t تم رسم دالة الارتباط الذاتي $r(k)$ للسلسلة كما هو موضح في الشكل (2) . وبفحص هذه الدالة بدقة وجد أن معاملات الارتباطات الذاتية قد تناقصت أولاً ولكنها مع مرور الزمن بدأت في الإرتفاع مرة أخرى ولذلك كان استقرار السلسلة محل شك كبير وبالتالي كان لزاماً تحويل البيانات الى بيانات أخرى مستقرة .

وقد استخدمت اللوغاريتمات لتحويل البيانات الأصلية y_t الى سلسلة أخرى y_t^* ثم أخذت الفروق الأولى للوغاريتمات y_t^* . ولقد استخدمت التحويلة اللوغاريتمية وذلك بغرض تسكين السلسلة خاصة تباينها .

إذا رمزنا للسلسلة الأصلية للحجاج بالرمز y_t ولسلسلة اللوغاريتمات بالرمز y_t^* ولسلسلة الفروق بالرمز z_t فإن :

$$y_t^* = \log y_t \quad , \ t = 1, 2, \ldots, 25 \qquad (1)$$

$$\begin{aligned} z_t &= y_t^* - y_{t-1}^* \\ &= \log y_t - \log y_{t-1} \quad , \ t = 2, 3, \ldots, 25 \qquad (2) \end{aligned}$$

ويلاحظ اننا نفقد مشاهدة من المشاهدات z_t بسبب أخذ الفروق الأولى وذلك من أجل تسكين السلسلة الزمنية للحجاج . وبفحص الرسم البياني لسلسلة الفروق z_t يمكن القول بأن السلسلة z_t هي سلسلة مستقرة الى حد كبير ، أنظر شكل (3) . وللتأكد من استقرار السلسلة

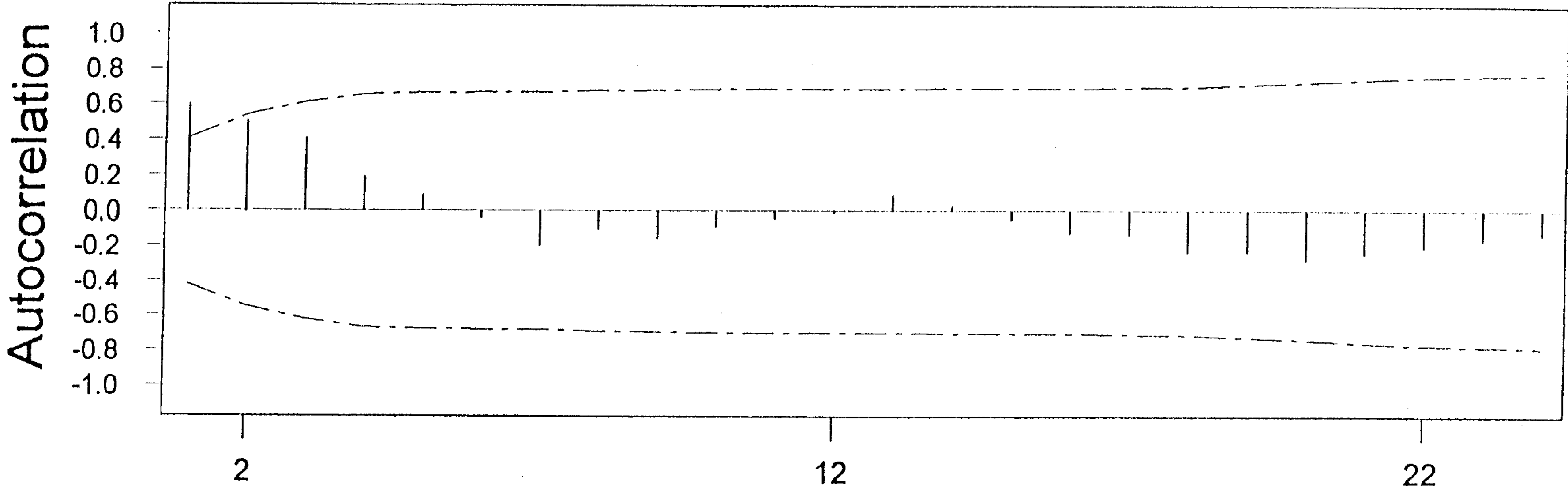


Fig.(2)

Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.60	2.99	10.09	8	-0.11	-0.32	26.35	15	-0.05	-0.14	28.52	22	-0.20	-0.55	71.75
2	0.51	1.94	17.64	9	-0.16	-0.47	27.38	16	-0.13	-0.37	29.70	23	-0.17	-0.45	81.22
3	0.41	1.38	22.86	10	-0.09	-0.27	27.75	17	-0.13	-0.39	31.16	24	-0.13	-0.35	92.75
4	0.20	0.61	24.09	11	-0.04	-0.13	27.84	18	-0.23	-0.68	36.43				
5	0.09	0.29	24.39	12	-0.01	-0.03	27.84	19	-0.23	-0.66	42.30				
6	-0.04	-0.13	24.45	13	0.09	0.27	28.31	20	-0.28	-0.78	52.51				
7	-0.20	-0.60	25.89	14	0.03	0.09	28.37	21	-0.24	-0.67	62.55				

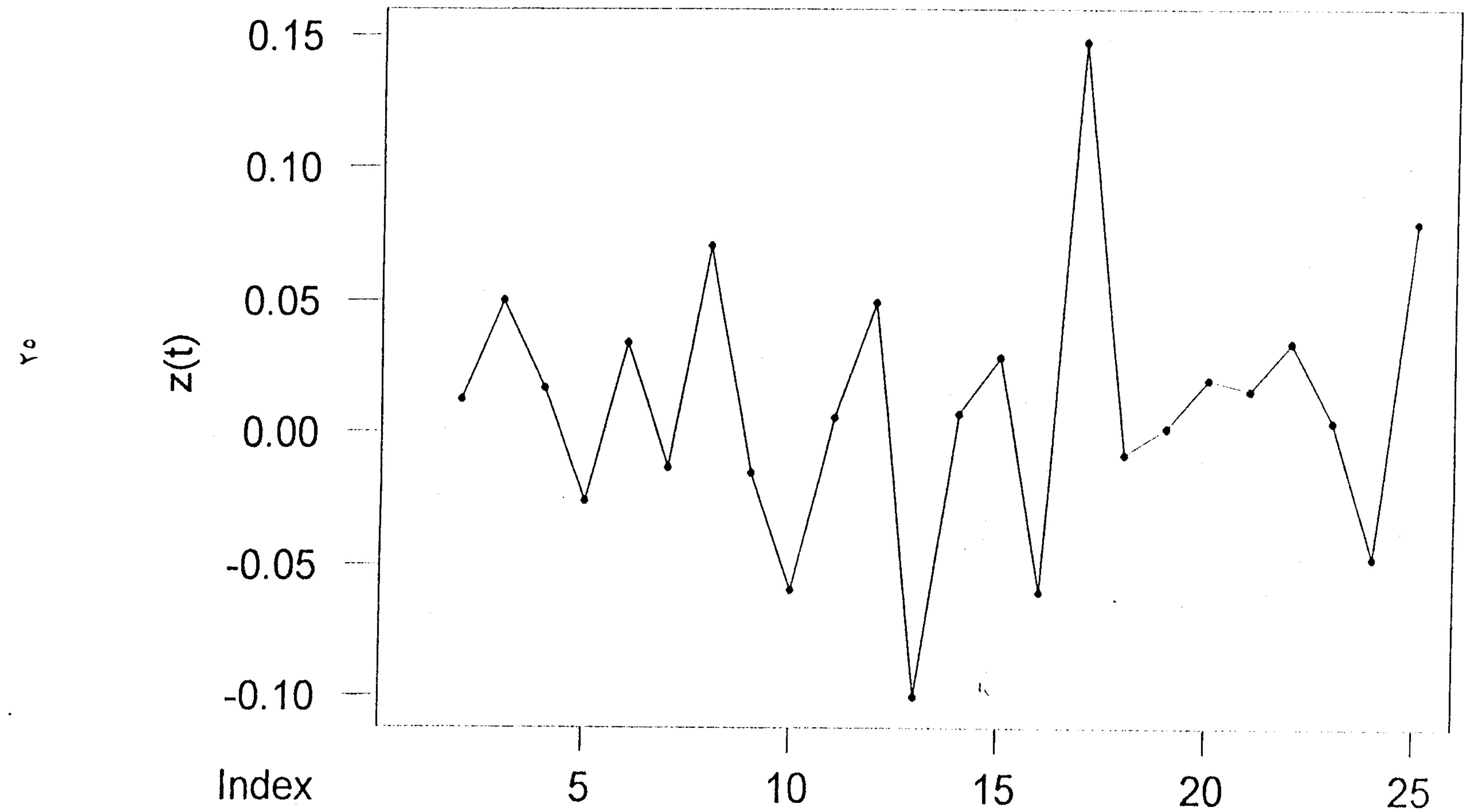

Time Series Plot Of The Difference Of Log [y(t)]
z(t)
0.15
0.10
0.05
0.00
-0.05
-0.10
Index
5
10
15
20
25

Fig.(3)

z_t المحولة تم حساب دالة الارتباط الذاتي لها ورسمها في شكل (4) والتي تميزت بتناقص واضح وسريع الى الصفر ، ومن ثم يمكن القول أن السلسلة z_t مستقرة وأن التحويلة المستخدمة قد نجحت في تحويل سلسلة الحجاج y_t غير المستقرة الى السلسلة z_t المستقرة بشكل يمكن معه استخدام الاسلوب الحديث للسلاسل الزمنية لنمذجة وتحليل هذه البيانات .

2. اختيار النموذج المبدئي المناسب (تعريف النموذج)

بمجرد التحقق من إستقرار السلسلة z_t ، فإن الخطوة التالية في إسلوب بوكس وجينكنز هي التعرف على النموذج المبدئي الملائم لبيانات هذه السلسلة وما إذا كان نموذج إنحدار ذاتي AR(P) أو متوسطات متحركة MA(q) أو نموذج مختلط ARMA (p,q) . والأدوات الهامة لهذه المرحلة من التحليل هي :

(أ) فحص دالة الارتباط الذاتي لبيانات السلسلة (العينة) ، ومقارنة هذه الدالة مع دالـــــة الارتباط الذاتي المناظرة في المجتمع .

(ب) فحص دالة الارتباط الذاتي الجزئي لبيانات السلسلة ومقارنة هذه الدالة مع دالة الارتباط الذاتي الجزئي المناظرة للمجتمع .

(أ) فحص دالة الارتباط الذاتي للسلسلة

قد يعتقد القارئ العادي بأن جميع معاملات الارتباط الذاتية لا تختلف معنوياً عن الصفر كما يتضح من رسم هذه الدالة (شكل 4) وذلك لأن جميع المعاملات كما تبدو (بغض النظر عن الاشارة) أقل من ضعف الخطأ المعياري . أي أن :

$$|r(k)| < 2\ SE\ (r(k)) \qquad ; \qquad k = 1,2,\ldots$$

ولكن الفحص الدقيق وخبرة الباحث المتمرس في التعامل مع هذه النوعية من البيانات لا يمكن أن يقبل مثل هذه الاستدلالات بالنسبة لفترات الزمن الابتدائية أي بالنسبة لـ $t = 1,2$.

Autocorrelation Function for z(t)

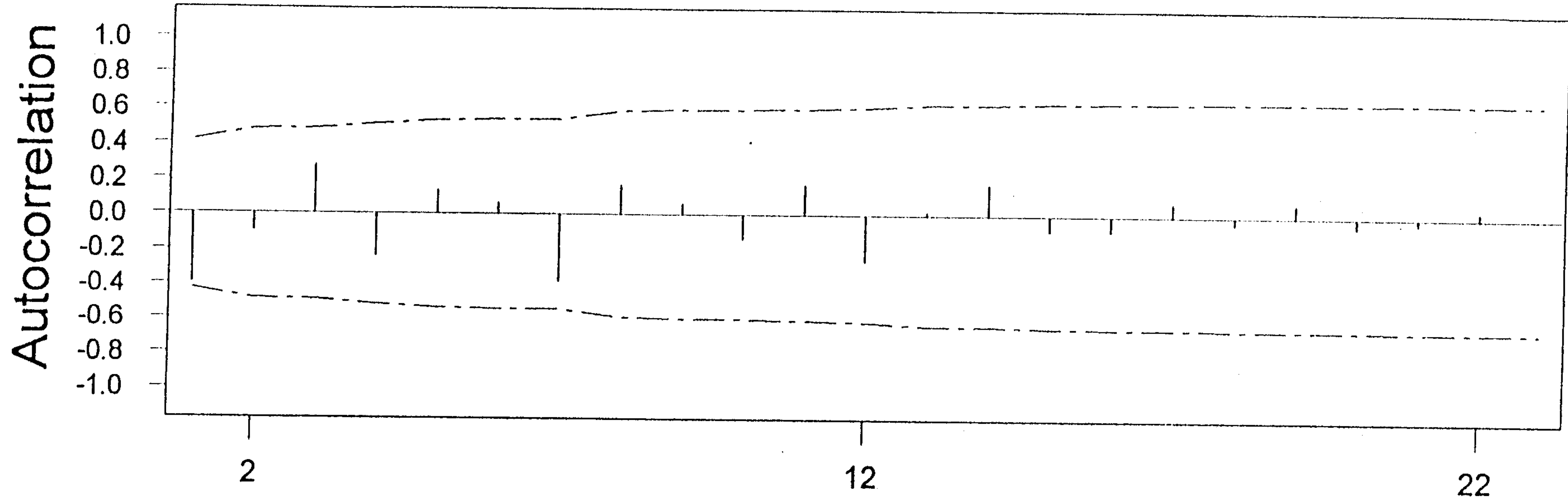


Fig.(4)

Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.39	-1.90	4.07	8	0.17	0.59	15.59	15	-0.08	-0.27	24.32	22	0.04	0.14	27.54
2	-0.10	-0.42	4.34	9	0.07	0.24	15.78	16	-0.08	-0.27	24.86	23	0.00	0.01	27.54
3	0.27	1.17	6.57	10	-0.14	-0.48	16.64	17	0.08	0.26	25.43				
4	-0.24	-0.96	8.35	11	0.18	0.60	18.11	18	-0.04	-0.12	25.59				
5	0.13	0.52	8.94	12	-0.26	-0.89	21.74	19	0.08	0.25	26.33				
6	0.07	0.26	9.10	13	0.02	0.08	21.77	20	-0.05	-0.16	26.73				
7	-0.38	-1.47	14.49	14	0.18	0.60	23.83	21	-0.03	-0.10	26.93				

ولذلك لا يمكن بسهولة قبول منطق أن الارتباط الذاتي من الرتبة الأولى $\rho(1)$ لا يختلف معنوياً عن الصفر ولذلك كان لابد من دراسة معاملي الارتباط الذاتي من الرتبة الأولى $\rho(1)$ والارتباط الذاتي من الرتبة الثانية $\rho(2)$ بمزيد من التدقيق .

بالنسبة لمعامل الارتباط الذاتي من الرتبة الأولى $\rho(1)$ نجد أن تقدير هذا المعامل هو $r(1) = -0.39$. ولإختبار معنوية $\rho(1)$ في المجتمع يكون لدينا الاختبار الإحصائي الأتي :

$$H_0 : \rho(1) = 0$$

$$H_1 : \rho(1) \neq 0$$

ولإجراء مثل هذا الاختبار يستخدم عادة الإحصاء T والذي يمكن تقريبه باستخدام التوزيع الطبيعي القياسي . ويأخذ هذا الاحصاء الصورة الأتية :

$$T = \frac{r(1)}{SE\ (r(1))}$$

حيث يمثل الرمز $SE\ (r(1))$ الخطأ المعياري للمقدر $r(1)$. ومن نتائج الحاسب الآلي (شكل 4) نجد أن قيمة هذا الاحصاء هو :

$$T = -1.90$$

وبإستخدام جداول التوزيع الطبيعي القياسي لإيجاد قيمة P (P-value) المناظرة نجد أن :

$$\frac{P}{2} = 0.0287$$

وبالتالي فإن :

$$P = 0.0574$$

وهذا يعني أن معامل الارتباط الذاتي من الرتبة الأولى $\rho(1)$ يختلف معنوياً عن الصفر إذا كان مستوى المعنوية أكبر من أو يساوي 0.06 . وبناءاً على ذلك نرفض فرض العدم H_0 ونستدل على أن $\rho(1)$ يختلف معنوياً عن الصفر .

وبفحص باقي معاملات الارتباط الذاتي يمكن القول بأن معامل الارتباط الذاتي من الرتبة الثانية لا يختلف معنوياً عن الصفر . ولكن بعد الفجوة الزمنية الثانية نجد أن بعض المعاملات الأقل أهمية تأخذ قيماً كبيرة بعض الشئ - وإن ظلت لا تختلف معنوياً عن الصفر - مثل قيم $r(3)$ و $r(7)$ ثم تتناقص دالة الارتباط بعد ذلك بشكل تدريجي . وبذلك يمكن الاستدلال على أن سلوك دالة الارتباط الذاتي $\rho(k)$ المناظرة للعملية العشوائية التي ولدت بيانات هذه السلسلة تأخذ أحد الشكلين الأتيين :

1. أنها تنقطع بشكل مفاجئ بعد الفجوة الزمنية الأولى وذلك إذا اعتبرنا أن جميع معاملات الارتباط الذاتي $\rho(k)\ ;\ k \geq 2$ لا تختلف معنوياً عن الصفر . وبذلك فإن هذه الدالة تشابه نظيرتها في نموذج المتوسطات المتحركة من الرتبة الأولى أي النموذج $MA(1)$.

2. أنها تتناقص بشكل تدريجي وذلك إذا اعتبرنا أن بعض معاملات الارتباط الذاتي - بالرغم من عدم معنوياتها - ذات قيم كبيرة نسبياً بشكل لا يمكن أهماله . وبذلك تشابه هذه الدالة نظيرتها في نماذج الانحدار الذاتي أو النماذج المختلطة .

وقبل اتخاذ قرار نهائي بشأن نمط دالة الارتباط الذاتي $\rho(k)$ يجب دراسة نمط معاملات الارتباط الذاتي الجزئي ϕ_{kk} للسلسلة أولاً .

(ب) فحص دالة الارتباط الذاتي الجزئي للسلسلة

بفحص دالة الارتباط الذاتي الجزئي للسلسلة أي ϕ_{kk} كما تظهر في الشكل (5) يمكن الاستدلال على أن جميع معاملات الارتباط الذاتي الجزئي بعد الفجوة الزمنية الثانية لا تختلف معنوياً عن الصفر . أما بالنسبة لمعاملي الارتباط الجزئي ϕ_{11} و ϕ_{22} فإنهما يحتاجان الى مزيد من الفحص والدراسة .

مرة أخرى نلاحظ أن تقدير معامل الارتباط الذاتي الجزئي من الرتبة الأولى $\hat{\phi}_{11}$ يساوي 0.39 - وهو يساوي معامل الارتباط الذاتي $r(1)$. وقيمة الاحصاء T المناظرة تساوي 1.90-

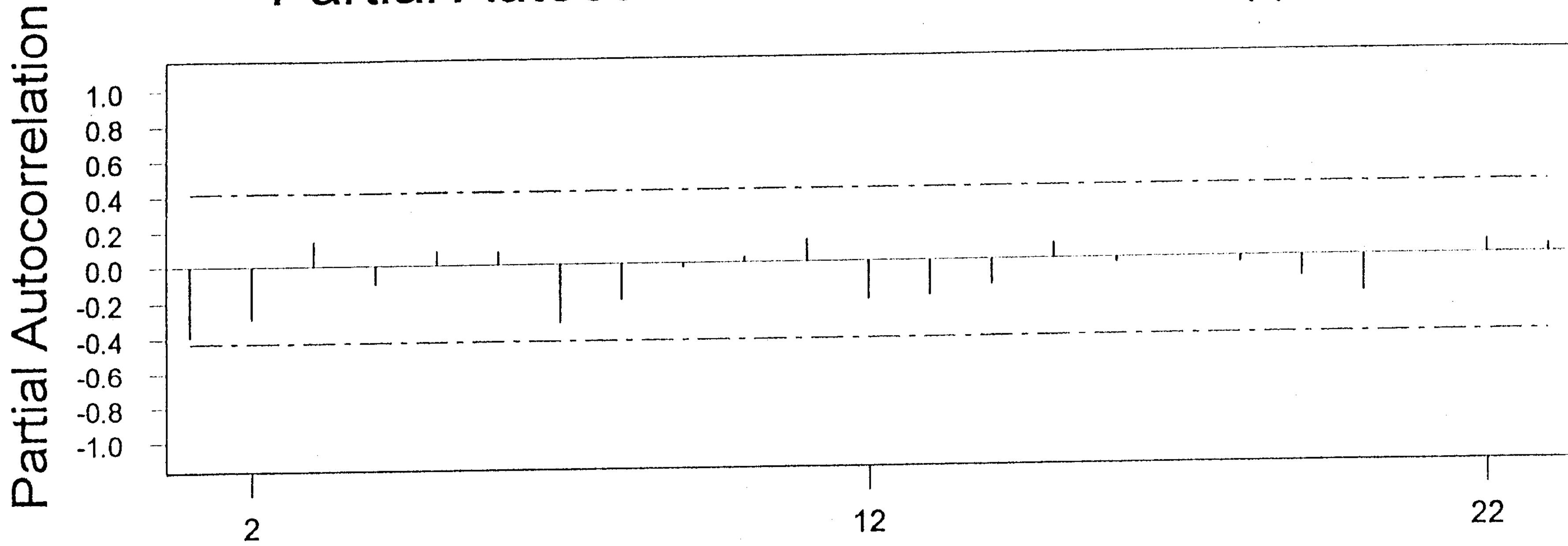


Fig.(5)

Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	-0.39	-1.90	8	-0.20	-1.00	15	0.09	0.44	22	0.08	0.39
2	-0.29	-1.42	9	-0.02	-0.12	16	-0.03	-0.14	23	0.05	0.25
3	0.15	0.71	10	0.04	0.18	17	-0.00	-0.01			
4	-0.10	-0.50	11	0.13	0.63	18	-0.04	-0.18			
5	0.09	0.42	12	-0.22	-1.06	19	-0.12	-0.58			
6	0.08	0.39	13	-0.20	-0.96	20	-0.21	-1.01			
7	-0.32	-1.58	14	-0.14	-0.70	21	-0.00	-0.00			

وبالتالي فإن قيمة P المناظرة تساوي 0.0574 ، ومن ثم نرفض فرض العدم ونستدل على أن قيمة ϕ_{11} في المجتمع تختلف معنوياً عن الصفر . أما بالنسبة لتقدير معامل الارتباط الذاتي الجزئي من الرتبة الثانية $\hat{\phi}_{22}$ فيساوي 0.29 - وأن قيمة الاحصاء T المناظرة تساوي 1.42- ومن ثم فإن قيمة P هي :

$$\frac{P}{2} = 0.077 \quad \Rightarrow \quad P = 0.1556$$

وبالتالي فإن معامل الارتباط في المجتمع ϕ_{22} لا يختلف معنوياً عن الصفر إذا كان مستوى المعنوية α أقل من 0.15 . أما إذا كان مستوى المعنوية المستخدم أكبر من أو يساوي 0.16 ، فإنه يمكن إعتبار أن قيمة ϕ_{22} في المجتمع تختلف معنوياً عن الصفر .

وبذلك يمكن الاستدلال على أن سلوك دالة الارتباط الجزئي ϕ_{kk} المناظرة للعملية العشوائية التي ولدت بيانات هذه السلسلة تأخذ أحد الأشكال الثلاثة الآتية :

1. أنها تنقطع تماماً بعد الفجوة الزمنية الأولى وذلك إذا إعتبرنا أن جميع معاملات الارتباط الذاتي الجزئي ϕ_{kk} , $k \geq 2$ لا تختلف معنوياً عن الصفر . وهذا السلوك يمكن قبوله بمستوى معنوية α يتراوح بين (0.06, 0.15) . وبذلك فإن هذه الدالة تشابه نظيرتها في نموذج الانحدار الذاتي من الرتبة الأولى أي النموذج AR(1) .

2. أنها تنقطع تماماًبعد الفجوة الزمنية الثانية وذلك إذا إعتبرنا أن ϕ_{22} تختلف معنوياً عن الصفر . وهذا السلوك يمكن قبوله إذا كان مستوى المعنوية المستخدم أكبر من أو يساوي 0.16 وبذلك فإن هذه الدالة تشابه نظيرتها في نموذج الانحدار الذاتي من الرتبة الثانية أي النموذج AR(2) .

3. أنها تتناقص بشكل تدريجي بدلاً من الافتراض أنها تساوي الصفر فجأة وذلك إذا اعتبرنا أن معاملات الارتباط الذاتي الجزئي - بالرغم من عدم معنوياتها - ذات أهمية معينة عند الفجوة الزمنية الثانية والفجوات الزمنية التالية .

بعد فحص دالتي الارتباط الذاتي والذاتي الجزئي نقترح ثلاثة نماذج مختلفة تستحق المزيد من الفحص والدراسة والمقارنة بينهم ، ومن المعتاد في التحليل الحديث للسلاسل الزمنية أن يكون لدينا أكثر من نموذج ملائم بشكل مبدئي للبيانات في مرحلة التعرف لأن الهدف الرئيسي لهذه المرحلة هو تضييق نطاق النماذج التي يمكن اختيارها لاجراء المزيد من الدراسة . والنماذج المختارة في هذه المرحلة هي :

(1) نموذج الانحدار الذاتي من الرتبة الأولى أو نموذج AR(1) . وقد اقترحنا هذا النموذج بإعتبار أن الدالة ϕ_{kk} تنقطع بعد الفجوة الزمنية الأولى وأن الدالة $\rho(k)$ تتناقص بشكل تدريجي .

(2) نموذج الانحدار الذاتي من الرتبة الثانية أو نموذج AR(2) . وقد إقترحنا هذا النموذج بإعتبار أن الدالة ϕ_{kk} تنقطع بعد الفجوة الزمنية الثانية وأن الدالة $\rho(k)$ تتناقص بشكل تدريجي .

(3) نموذج المتوسطات المتحركة من الرتبة الأولى أو نموذج MA(1) . وقد إقترحنا هذا النموذج بإعتبار أن الدالة $\rho(k)$ تنقطع بعد الفجوة الزمنية الأولى وأن الدالة ϕ_{kk} تتناقص بشكل تدريجي .

وهذا يعني أن النماذج الثلاثة التي نرشحها لمزيد من الدراسة والفحص لتحليل سلسلة الحج الأصلية y_t هي :

النموذج الأول (I)

$$(1 - \phi_1 B)\, \Delta \log y_t = \delta + \varepsilon_t$$

النموذج الثاني (II)

$$(1 - \phi_1 B - \phi_2 B^2)\, \Delta \log y_t = \delta + \varepsilon_t$$

النموذج الثالث (III)

$$\Delta \log y_t = \delta + (1 - \theta_1 B)\, \varepsilon_t$$

ونقوم في الأجزاء التالية بإجراء فحص مفصل لكل نموذج من هذه النماذج :

2. تقديـــــــر النماذج

نمــــوذج (I)

إستخدمت طريقة المربعات الصغرى غير الخطية لتقدير المعالم δ ,ϕ_1 باستخدام برنامج MINITAB-VERSION 12 . ويحتوي جدول (1) على التقديرات والأخطاء المعياريـــــــــة للمقدرات بالإضافة الى قيم الاحصاء T التي تستخدم لإختبار معنوية المعالم . وقد حسبنا قيمة P (P-Value) المناظرة للمقدر $\hat{\delta}$ فكانت 0.1738 ، أما قيمة P المناظرة للمقدر $\hat{\phi}_1$ فكانت 0.0384 .

نمــــوذج (II)

يتضمن جدول (2) التقديرات التي حصلنا عليها باستخدام الحاسب الآلي والأخطاء المعيارية وقيم الإحصاء T لمقدرات النموذج (II) . وقد حسبنا قيم P للمقدرات الثلاثة $\hat{\delta}$, $\hat{\phi}_1$, $\hat{\phi}_2$. فكانت على الترتيب 0.0802 , 0.0114 , 0.1362 .

نمــــوذج (III)

يتضمن جدول (3) التقديرات التي حصلنا عليها باستخدام الحاسب الآلي والأخطاء المعيارية للمقدرات وقيم الإحصاء T لمقدرات النموذج (III) . وقد حسبنا قيم P المناظرة للمقدر $\hat{\delta}$ فكانت 0.069 ، أما قيمة P المناظرة للمقدر $\hat{\theta}_1$ فكانت 0.008 .

3. فحص النماذج المقدرة وتشخيصها

المرحلة الثالثة من مراحل تطبيق أسلوب بوكس وجينكنز هي تشخيص النماذج المقدرة لمعرفة مدى ملاءمة كل نموذج وكفاءته وإمكانية تحسينه أو تطويره . ولقد أجريت ثلاث فحوص رئيسية لفحص وتقويم كل نموذج على حدة وهي :

Table (1)

Estimation of the First Model

ARIMA Model

ARIMA model for Log(y)

Estimates at each iteration

Iteration	SSE	Parameters	
0	0.260491	0.100	0.099
1	0.109383	-0.050	0.057
2	0.065355	-0.200	0.034
3	0.051318	-0.350	0.018
4	0.050479	-0.418	0.013
5	0.050478	-0.422	0.013
6	0.050478	-0.422	0.013

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type	Coef	StDev	T
AR 1	-0.4218	0.2035	-2.07
Constant	0.013315	0.009786	1.36

Differencing: 1 regular difference
Number of observations: Original series 25, after differencing 24
Residuals: SS = 0.0504770 (backforecasts excluded)
MS = 0.0022944 DF = 22

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	13.1 (DF=11)	* (DF= *)	* (DF= *)	* (DF= *)

Table (2)

Estimation of the Second Model

ARIMA Model

```
ARIMA model for Log(y)

Estimates at each iteration
Iteration       SSE      Parameters
    0        0.223271     0.100     0.100     0.088
    1        0.087542    -0.050     0.075     0.044
    2        0.059557    -0.200    -0.032     0.029
    3        0.049181    -0.350    -0.152     0.021
    4        0.045998    -0.500    -0.281     0.016
    5        0.045886    -0.533    -0.318     0.017
    6        0.045886    -0.535    -0.330     0.017
    7        0.045884    -0.537    -0.325     0.017
Unable to reduce sum of squares any further

Final Estimates of Parameters
Type            Coef        StDev          T
AR   1       -0.5371       0.2124      -2.53
AR   2       -0.3253       0.2183      -1.49
Constant   0.016670     0.009534       1.75

Differencing: 1 regular difference
Number of observations:  Original series 25, after differencing 24
Residuals:       SS = 0.0457112   (backforecasts excluded)
                 MS = 0.0021767   DF = 21

Modified Box-Pierce (Ljung-Box) Chi-Square statistic
Lag                   12             24              36              48
Chi-Square    11.3(DF=10)      * (DF= *)       * (DF= *)       * (DF= *)
```

Table (3)

Estimation of the Third Model

ARIMA Model

```
ARIMA model for Log(y)

Estimates at each iteration
Iteration        SSE      Parameters
    0        0.351346    0.100     0.110
    1        0.062883    0.250     0.026
    2        0.049474    0.400     0.013
    3        0.048080    0.491     0.009
    4        0.048075    0.496     0.009
    5        0.048075    0.496     0.009
Relative change in each estimate less than  0.0010

Final Estimates of Parameters
Type            Coef        StDev          T
MA   1        0.4960       0.1866       2.66
Constant  0.008858     0.004855       1.82

Differencing: 1 regular difference
Number of observations:  Original series 25, after differencing 24
Residuals:      SS = 0.0479470  (backforecasts excluded)
                MS = 0.0021794  DF = 22

Modified Box-Pierce (Ljung-Box) Chi-Square statistic
Lag                    12             24             36             48
Chi-Square    12.5(DF=11)      * (DF= *)      * (DF= *)      * (DF= *)
```

A. تحليل البواقي Residual Analysis

B. إضافة معالم أخرى الى النموذج Overfitting

C. حذف بعض معاملات النموذج Underfitting

ونقوم فيما يلي بعرض نتائج هذه الفحوص التشخيصية بشئ من التفصيل لكـــل نموذج على حدة :

تشخيص نموذج (I)

A. تحليل البواقـــي

ويقصد بتحليل البواقي مجموعة من الاختبارات والفحوص التشخيصية تجرى على البواقي الناتجة من توفيق النموذج لتقويم كفاءة النموذج ومدى ملاءمة الفروض النظرية التي يقوم عليها . وتضم هذه المجموعة من الاختبارات الفحوص الأتية :

1. فحص دالة الارتباط الذاتي للبواقي .
2. فحص دالة الارتباط الذاتي الجزئي للبواقي
3. إختبار Q (بوكس وبيرس المعدل)
4. فحص نموذج الفروق الأولى للبواقي ($\Delta\ e_t$)
5. رسم البواقي كسلسلة زمنية
6. رسم البواقي في مقابل القيم المقدرة
7. فحص وإختبار إعتيادية الأخطاء

ونعرض فيما يلي أهم نتائج هذه الفحوص التشخيصية والإختبارات الخاصة بنموذج (I) .

1. فحص دالة الارتباط الذاتي للبواقي

إذا كان النموذج (I) ملائماً ، فإن البواقي الناجمة من توفيق هذا النموذج للبيانات يجب أن تكون تغيرات عشوائية بحتة . وبفحص دالة الارتباط الذاتي للبواقي (شكل 6) نجد أن كل معامل ارتباط ذاتي يقع داخل فترة الثقة للعينات الكبيرة 95% ، أي أن هذه الدالة خالية مـــن

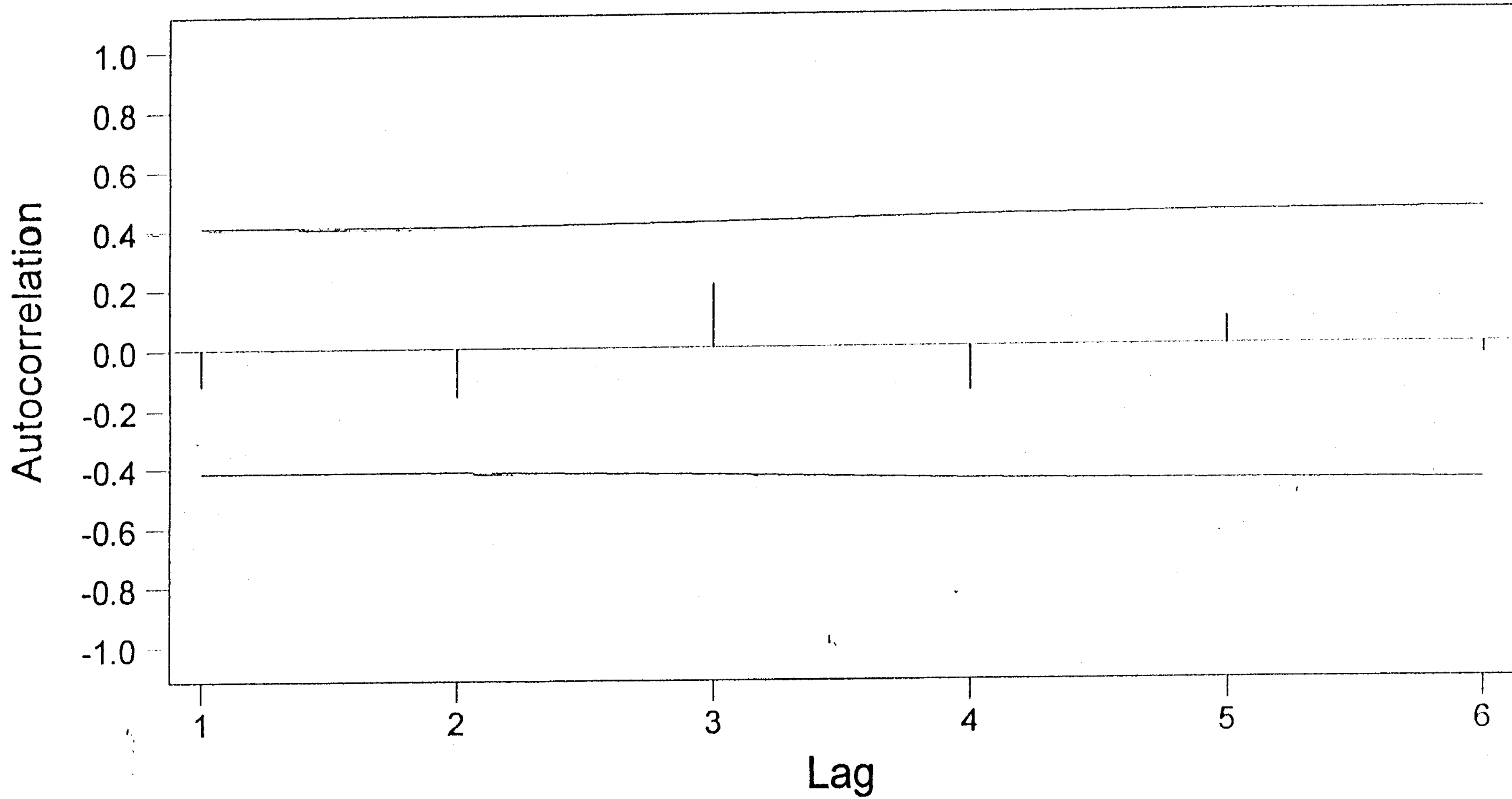


Fig.(6)

ينتقل إلى ص 38

أي نتوءات . ولذلك يمكن الاستدلال على أن كل معامل من معاملات الارتباط الذاتي في المجتمع لا يختلف معنوياً عن الصفر ، وهذا مؤشر جيد على أن البواقي هي تغيرات عشرائية بحتة Pure random . وتجدر الاشارة هنا الى أن جميع الاختبارات والفحوص التي تجري هنا تنفذ بشكل منفرد ومستقل ومستوى معنوية خاص بكل إختبار على حدة .

2. فحص دالة الارتباط الذاتي الجزئي للبواقي

بفحص شكل (7) يمكن الاستدلال على أن دالة الارتباط الذاتي الجزئي لا تظهر أي إختلال في فرض عشوائية الأخطاء الحقيقية ε_t ، وهذا مؤشر أخر على ملائمة أحد الفروض الهامة التي يقوم عليها النموذج (I) .

3. إختبار Q المعدل (Ljung - Box)

بدلاً من فحص وإختبار كل معامل إرتباط ذاتي على حدة - كما في الإختبارين السابقين - قد يكون من الملاءم أحياناً معرفة ما إذا كانت مجموعة عددها k من معاملات الارتباط الذاتي تدل على ملاءمة النموذج المختار . ولقد قدم العالمان بوكس وبيرس Box - Pierce الإحصاء Q لأول مرة عام 1970 من أجل تحقيق هذا الهدف في الصورة الأتية :

$$Q = n \sum_{i=1}^{k} r_e^2\ (i)$$

ولقد أثبتا هذان العالمان أن الاحصاء Q يتبع بشكل تقريبي توزيع χ^2 بدرجات حرية (k-p-q) وذلك بافتراض أن الأخطاء ε_t تمثل تغيرات عشوائية بحتة ، أي بافتراض صحة H_0 المعرف على الصورة الأتية :

$$H_0 : \rho_\varepsilon(1) = \rho_\varepsilon(2) = \ldots = \rho_\varepsilon(k) = 0$$

ولقد إقترحا العالمان Ljung - Box عام 1978 تعديل هذا الاحصاء من أجل تحسين الأداء في حالة السلاسل الزمنية القصيرة ليصبح على الصورة الأتية :

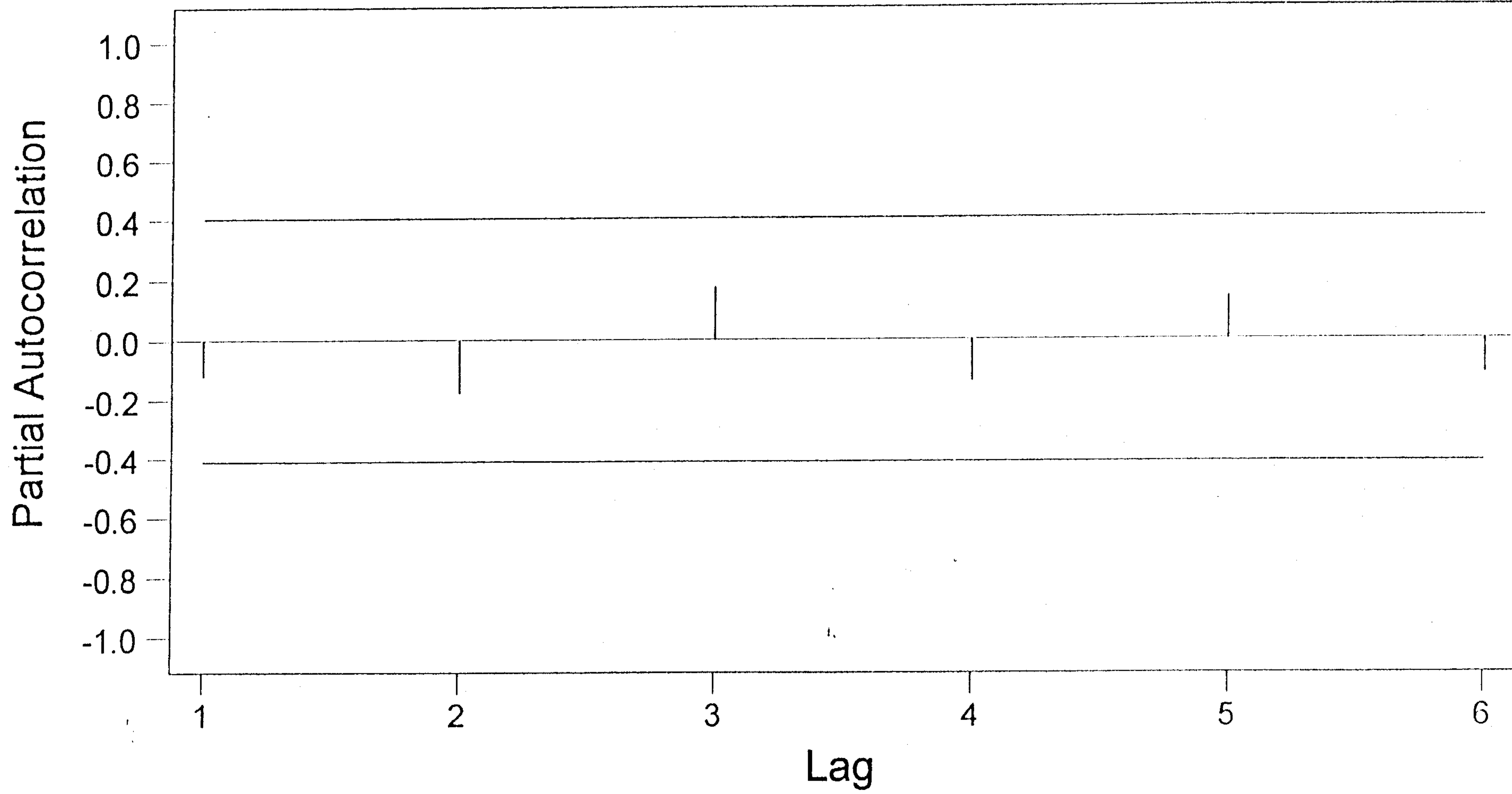

PACF of Residuals for Log(y)
(with 95% confidence limits for the partial autocorrelations)
Partial Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
1
2
3
4
5
6
Lag

Fig.(7)

$$Q = n(n+2) \sum_{i=1}^{k} \frac{r_e^2(i)}{n-k}$$

ولقد حسبت قيمة الإحصاء Q المعدل للنموذج (I) باستخدام الحاسب الآلي لأول 12 معامل إرتباط ذاتي $(k = 12)$ فكانت

$$Q(12) = 13.1 \qquad \ldots (1)$$

وبإستخدام جداول χ^2 بدرجات حرية 11 وباستخدام مستوى المعنوية $\alpha = 0.05$ كانت قيمة χ^2 الجدولية التي تحصر على يسارها مساحة مقدارها 0.95 هي

$$\chi^2_{0.95,11} = 19.7 \qquad \ldots (2)$$

من (1) ، (2) نجد أن قيمة الاحصاء Q المحسوبة أقل كثيراً من قيمة χ^2 الجدولية ، ومن ثم يستدل على عدم وجود نمط غير عشوائي في أول 12 معامل إرتباط ذاتي . ويعتبر هذا مؤشراً هاماً أخر على ملاءمة النموذج (I) لتحليل سلسلة الحج .

4. فحص الفروق الأولى للبواقي

إذا كانت الأخطاء الحقيقية ε_t تتبع عملية تغيرات عشوائية بحتة ، فإن الفروق الأولى للبواقي أي Δe_t يجب أن تتبع عملية MA(1) بمعلمة $\theta_1 = 1$. كما أن معامل الارتباط الذاتي من الرتبة الأولى يجب أن يساوي 0.5 - . وبفحص دالتي الارتباط الذاتي والذاتي الجزئي (شكل 8 وشكل 9) يمكن الاستدلال بسهولة على أن النموذج الملاءم لفروق البواقي Δe_t هو بالفعل نموذج MA(1) . ولقد قمنا بتوفيق هذا النموذج للبيانات Δe_t فوجدنا أن تقدير θ_1 هو $\hat{\theta}_1 = 0.923$ (أنظر جدول 4) وأن تقدير معامل الارتباط الذاتي الأول هو $r(1) = -.45$. ولذلك كان لزاماً علينا إجراء الإختبارين الأتيين :

$$\text{(i) } H_o: \theta_1 = 1 \quad \text{VS} \quad H_1: \theta_1 \neq 1$$

$$\text{(ii) } H_o: \rho(1) = -0.5 \quad \text{VS} \quad H_1: \theta \neq -0.5$$

ρ(1)

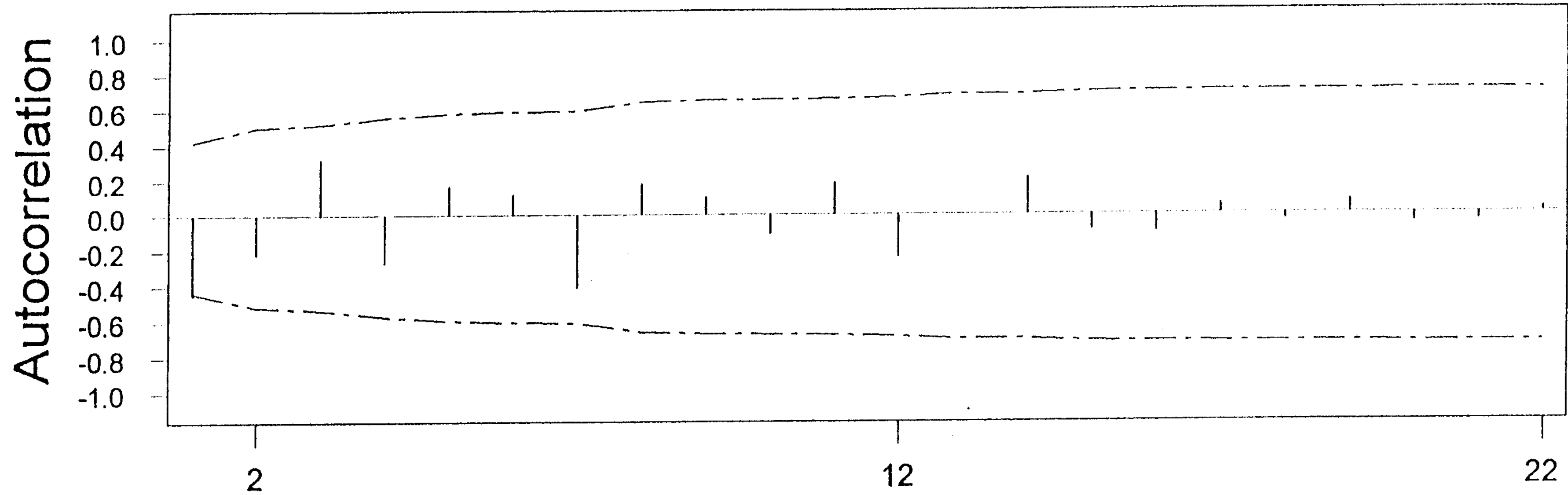


Fig.(8)

Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.45	-2.14	5.20	8	0.19	0.59	20.46	15	-0.09	-0.25	29.65	22	0.04	0.10	33.60
2	-0.22	-0.88	6.48	9	0.10	0.33	20.91	16	-0.10	-0.29	30.45				
3	0.33	1.28	9.53	10	-0.11	-0.33	21.40	17	0.06	0.18	30.81				
4	-0.27	-0.99	11.72	11	0.19	0.58	23.06	18	-0.03	-0.10	30.95				
5	0.17	0.61	12.66	12	-0.24	-0.74	26.08	19	0.08	0.24	31.90				
6	0.12	0.42	13.16	13	0.00	0.01	26.08	20	-0.05	-0.15	32.41				
7	-0.41	-1.41	19.14	14	0.22	0.65	29.11	21	-0.04	-0.12	32.88				

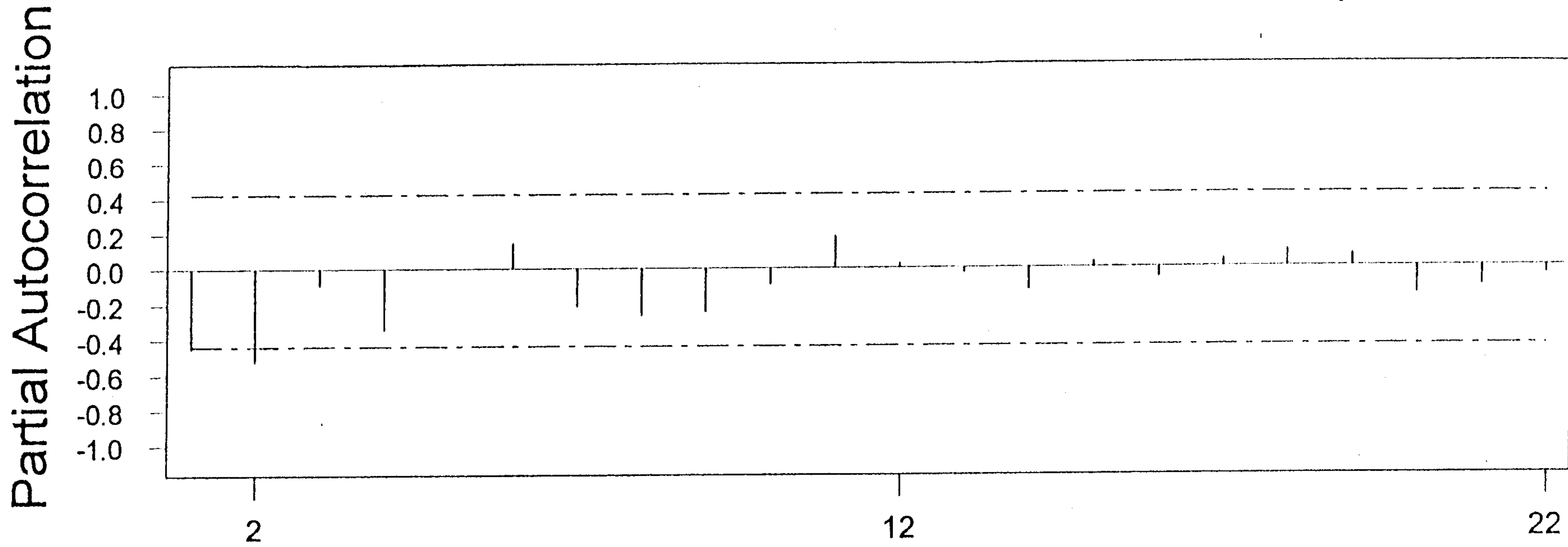


Fig.(9)

Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	-0.45	-2.14	8	-0.26	-1.25	15	0.03	0.16	22	-0.04	-0.21
2	-0.52	-2.48	9	-0.24	-1.16	16	-0.05	-0.26			
3	-0.09	-0.42	10	-0.09	-0.42	17	0.05	0.22			
4	-0.34	-1.63	11	0.19	0.89	18	0.10	0.48			
5	-0.00	-0.01	12	0.03	0.13	19	0.07	0.35			
6	0.15	0.71	13	-0.03	-0.13	20	-0.15	-0.73			
7	-0.21	-1.00	14	-0.12	-0.58	21	-0.11	-0.52			

Table (4)

Estimation For the Difference of the Residuals

ARIMA Model

ARIMA model for Diff.(Residuals)

Estimates at each iteration

Iteration	SSE	Parameters
0	0.101362	0.100
1	0.088955	0.250
2	0.078455	0.400
3	0.069760	0.550
4	0.062761	0.700
5	0.056956	0.850
6	0.054728	0.925
7	0.054726	0.924

Unable to reduce sum of squares any further

Final Estimates of Parameters

Type	Coef	StDev	T
MA 1	0.9236	0.1530	6.04

Number of observations: 23

Residuals: SS = 0.0545851 (backforecasts excluded)
MS = 0.0024811 DF = 22

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	13.3(DF=11)	* (DF= *)	* (DF= *)	* (DF= *)

ولإجراء الأختبار الأول نستخدم الاحصاء T على الصورة الأتية :

$$T = \frac{\hat{\theta}_1 - 1}{SE\ (\hat{\theta}_1)}$$

وباستخدام نتائج الحاسب الألي حسبنا قيمة هذا الاحصاء فوجدنا أنها :

$$T^* = \frac{0.923 - 1}{0.1772} = -0.43$$

وبمقارنة |T| بالقيمة الجدولية 2 والتي نحصل عليها من التوزيع الطبيعي القياسي نستدل على أن قيمة المعلمة θ_1 لا تختلف معنوياً عن 1 . ومن ثم يستدل على أن النموذج الملاءم للفروق الأولى للبواقي هو نموذج MA(1) بمعلمة تساوي 1 . وهذا دليل أخر على أن الأخطاء ε_t تتبع تغيرات عشوائية بحتة .

وتجدر الاشارة الى أننا لا نستطيع استخدام قيمة الاحصاء T المحسوبة في شكل (8) لاجراء الإختبار الثاني ، لأن قيمة هذا الاحصاء تستخدم فقط لدراسة معنوية إختلاف $\rho(1)$ عن الصفر وليس عن القيمة 0.5 - . ومن ثم كان لزاماً علينا أن نحسب قيمة الخطأ المعياري للمقدر $r(1)$ أولاً ثم نجري الاختيار المطلوب :

$$SE\ [r(1)] = \frac{r(1)}{T} = (-0.45)\ /\ (-2.14)$$

$$= 0.2103$$

ومن ثم فإن قيمة الاحصاء T لاجراء هذا الإختيار هي :

$$T = \frac{r(1) - \rho(1)}{SE\ [r(1)]}$$

$$= \frac{-0.45 + 0.5}{0.2103} = 0.2378$$

وبمقارنة هذه القيمة بالقيمة الجدولية 2 ، نستدل على أن معامل الإرتباط الذاتي الأول لسلسلة الفروق الأولى للأخطاء لا يختلف معنوياً عن 0.5 - . وهذا مؤشر أخر على أن الأخطاء ε_t تتبع تغيرات عشوائية بحتة .

5. رسم البواقي كسلسلة زمنية

بفحص هذا الرسم (شكل 10) يمكن القول بأن تشتت البواقي e_t حول خط الوسط صفر ثابت الى حد كبير مما يؤيد ملاءمة فرض ثبات تباين الأخطاء σ^2 .

6. رسم البواقي ضد القيم المقدرة

بفحص (شكل 11) يمكن القول بأنه يمكن تمثيله بشكل حزمة موازية للقيم المقدرة مما يؤيد القول بأنه ليس هناك ارتباط واضح بين البواقي e_t والقيم المقدرة ، وهذا يؤيد أحد الفروض الهامة النظرية في تحليل الانحدار والسلاسل الزمنية .

7. إختبار إعتيادية الأخطاء

لاختبار اعتيادية الأخطاء ε يمكن استخدام Normal probability plot . فشكل (12) يؤيد الفرض القائل بأن التغيرات العشوائية تتبع توزيع طبيعي . وللتحقيق من هذا حسب إحصاء اندرسون – دارلينج Anderson-Darling فوجد أن قيمة هذا الاحصاء تساوي 0.313 وأن قيمة P المناظرة تساوي 0.526 مما يؤيد فرض العدم القائل بأن الأخطاء تتبع توزيع طبيعي .

ولتلخيص ماسبق يمكن القول بأن النموذج (I) والذي سبق إختياره في مرحلة التعريف قد اجتاز بنجاح جميع الاختبارات المستخدمة في تحليل البواقي ، مما يزيد الثقة في كفاءة هذا النموذج واستخدامه في التنبؤ المستقبلي .

B . إضافة معالم أخرى الى النموذج Overfitting

لقد أضفنا الى نموذج (I) معلمة متوسطات متحركة من الرتبة الأولى . وباستخدام نتائج

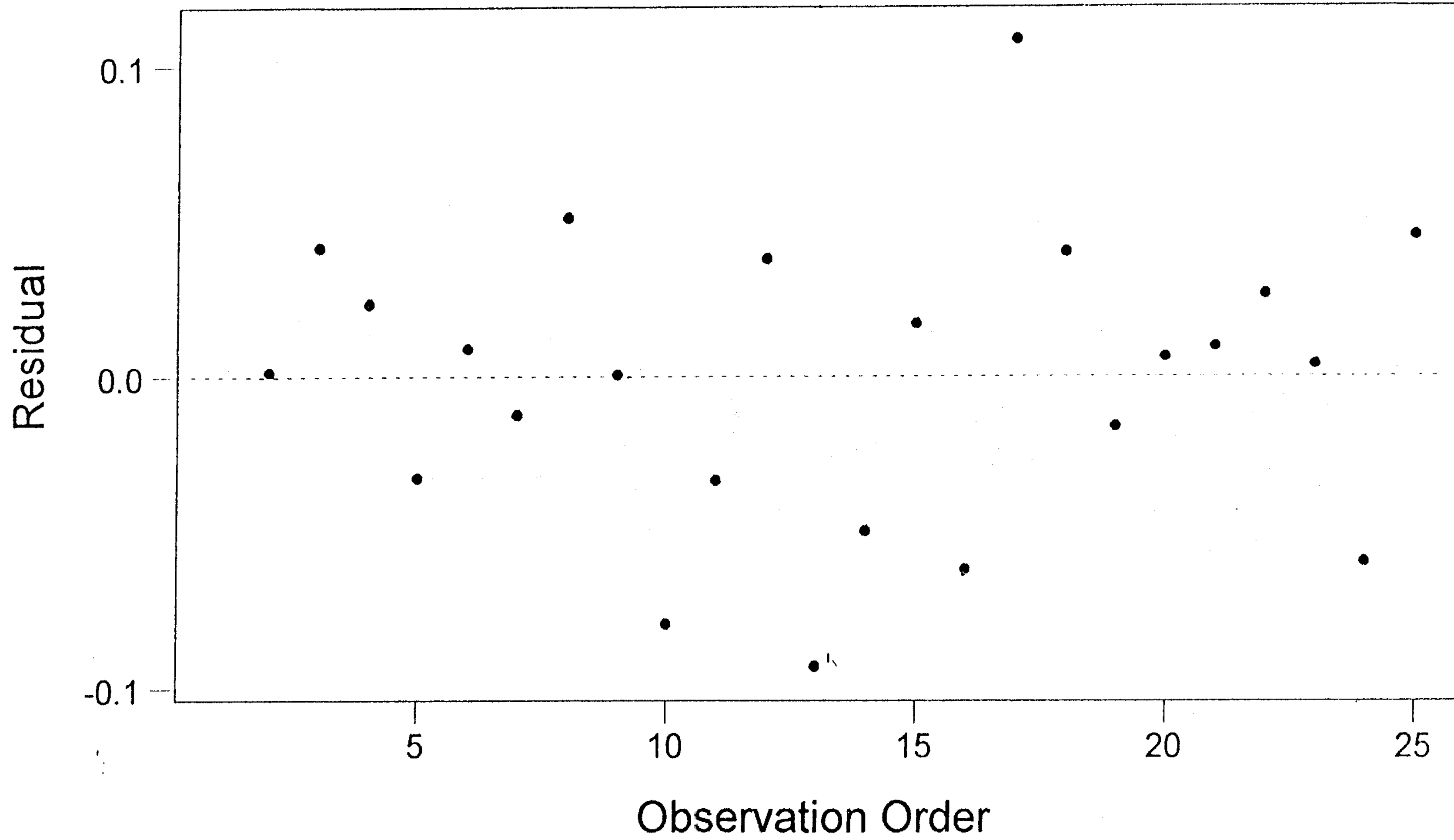

Residuals Versus the Order of the Data
(response is Log(y))
Residual
0.1
0.0
-0.1
5
10
15
20
25
Observation Order

Fig.(10)

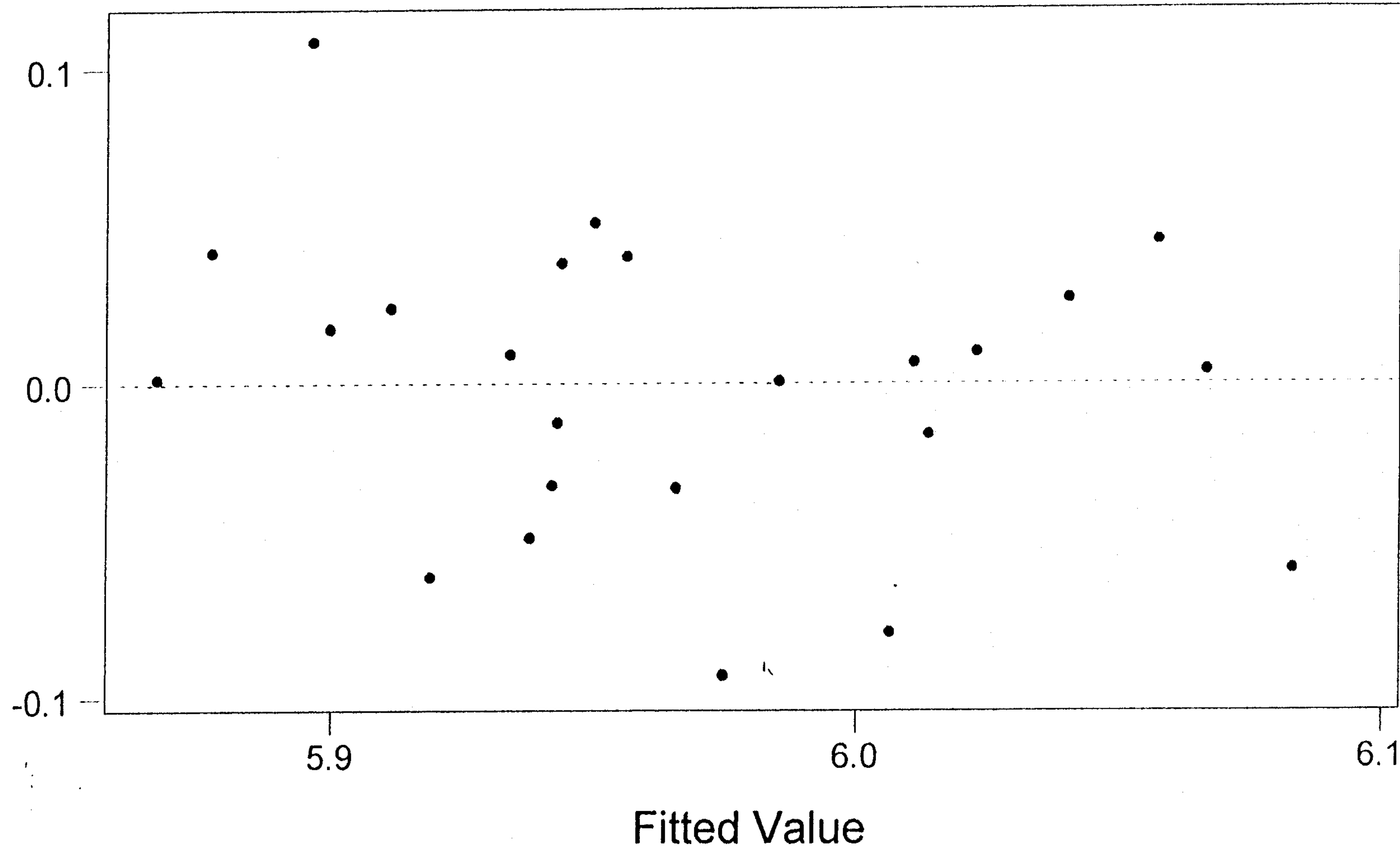

Residuals Versus the Fitted Values
(response is Log(y))
Residual
0.1
0.0
-0.1
5.9
6.0
6.1
Fitted Value

Fig.(11)

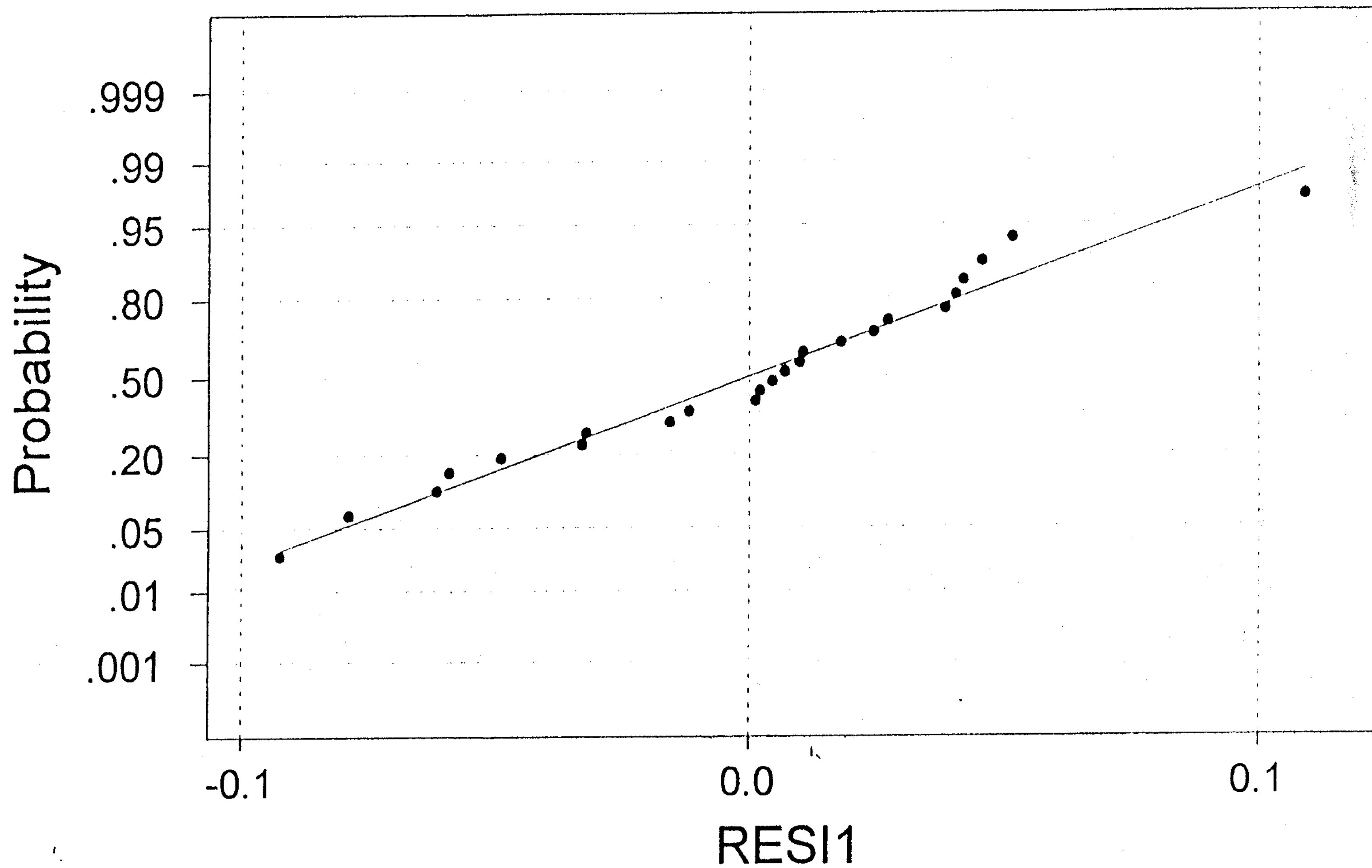

Normal Probability Plot
Probability
.999
.99
.95
.80
.50
.20
.05
.01
.001
-0.1
0.0
0.1
RESI1
Average: 0.0000271
StDev: 0.0468471
N: 24
Anderson-Darling Normality Test
A-Squared: 0.313
P-Value: 0.526

Fig.(12)

التقدير الموضحة في جدول (5) نجد أن قيمة مقدر معلمة المتوسطات المتحركة المضافة تساوي 0.4071 وأن قيمة إحصاء T المناظرة تساوي 0.99 ، مما يدل على أن معلمة المتوسطات المتحركة المضافة لا تختلف معنوياً عن الصفر . كما يلاحظ أن قيمة مقدر معلمة الانحدار الذاتي قد أصبحت 0.1339 - وقيمة إحصاء T المناظرة قد أصبحت 0.25 - ، أي أن معلمة الانحدار الذاتي قد أصبحت لا تختلف معنوياً عن الصفر بعد إضافة معلمة المتوسطات المتحركة من الرتبة الأولى ، وهذا يدل على شدة الارتباط بين مقدر معلمة المتوسطات المتحركة من الرتبة الأولى ومقدر معلمة الانحدار الذاتي من الرتبة الأولى ، وهذا دليل على أن وجود معلمة منهما يغني عن الأخرى .

ولقد أضفنا أيضاً معلمة انحدار ذاتي من الرتبة الثانية ، وهذا يؤدي الى نموذج (II) والذي سيتم تشخيصه في الجزء التالي

C . حذف بعض معاملات النموذج

نموذج (I) الذي إختير في مرحلة التعريف يحتوي على معامل واحد هو ϕ_1 ، وبالتالي فهو أصغر نموذج انحدار ذاتي ممكن ، ومن ثم فليس هناك ضرورة من حذف المعلمة خاصة أن قيمة الاحصاء T المناظرة تساوي 2.07 - مما يدل على اختلاف هذه المعلمة معنوياً عن الصفر.

من الواضح أن نتائج إضافة بعض المعالم وحذف بعض المعالم الأخرى تمدنا بمؤشرات إضافية - بجانب تحليل البواقي - على ملاءمة نموذج (I) لتحليل بيانات الحج . ومن ثم نستطيع القول في نهاية مرحلة التشخيص أن كل الوسائل والأدوات المستخدمة في هذه المرحلة تؤيد استخدام هذا النموذج لتحليل بيانات الحج . وبالرغم من شعورنا بالارتياح الى إستخدام هذا النموذج للتنبؤ بأعداد الحجاج في السنوات القادمة ،فقد رأينا أنه من الضروري فحص النموذجين الأخريين الذي تم ترشيحهما في مرحلة التعرف باستخدام نفس الطرق والوسائل التي

Table (5)

Estimation of ARIMA(1,1,1) Model
Over fitting

RIMA Model

```
RIMA model for Log(y)

stimates at each iteration
teration        SSE        Parameters
  0          0.300333     0.100     0.100     0.099
  1          0.078531    -0.043     0.243    -0.016
  2          0.058207     0.024     0.393    -0.003
  3          0.048279    -0.045     0.469     0.011
  4          0.047939    -0.114     0.406     0.010
  5          0.047938    -0.114     0.406     0.010
  6          0.047938    -0.114     0.407     0.010
  7          0.047938    -0.114     0.407     0.010
lative change in each estimate less than  0.0010

al Estimates of Parameters
pe          Coef         StDev         T
  1       -0.1139       0.4541     -0.25
  1        0.4071       0.4098      0.99
stant    0.009954     0.005793      1.72

ferencing: 1 regular difference
ber of observations:  Original series 25, after differencing 24
iduals:     SS = 0.0478443  (backforecasts excluded)
            MS = 0.0022783  DF = 21

fied Box-Pierce (Ljung-Box) Chi-Square statistic
                      12              24              36              48
Square    12.0(DF=10)       * (DF= *)       * (DF= *)       * (DF= *)
```

تشخيص النموذج (II)

سبق وأن ذكرنا أن النموذج (II) والذي يمكن صياغته كالتالي :

$$(1 - \phi_1 B - \phi_2 B^2)\, \Delta (\log y_t) = \delta + \varepsilon_t$$

يمثل أحد النماذج الثلاثة التي إخترناها في مرحلة التعريف لتحليل بيانات الحج . ولقد قمنا بتطبيق جميع الفحوص والإختبارات - التي إستخدمت لدراسة ملاءمة النموذج (I) - لتشخيص النموذج (II) . ويمكن تلخيص النتائج الأساسية لهذه الفحوص والإختبارات فيما يلي :

1. لا تحتوي دالة الارتباط الذاتي للفروق (أو دالة الارتباط الذاتي الجزئي) على أي نتوءات .
2. قيمة الإحصاء Q المعدل تساوي 11.3 .
3. الفروق الأولى للبواقي Δe_t لها نموذج MA(1) بمعلمة θ_1 لا تختلف معنوياً عن 1 ومعامل إرتباط ذاتي من الرتبة الأولى لا يختلف معنوياً عن 0.5- .
4. تشتت البواقي حول خط الوسط صفر ثابت تقريباً .
5. رسم البواقي ضد القيم المقدرة لا يظهر أي إرتباط واضح .
6. قيمة إحصاء اندرسون - دارلينج للبواقي تساوي 0.345 وقيمة P المناظرة تساوي 0.456.
7. إضافة معلمة متوسطات متحركة ليس له ضرورة .
8. إضافة معلمة انحدار ذاتي ثالثة ليس له ضرورة .

وبعد الفحص الدقيق للبواقي وإجراء الاختبارات الضرورية لهذه المرحلة أمكننا الاستدلال على ملاءمة هذا النموذج أيضاً لتحليل بيانات الحج .

ولقد شعرنا بأنه ليس هناك فروق جوهرية تذكر بين نتائج الفحوص التشخيصية لهذا النموذج ونتائج الفحوص التشخيصية للنموذج (I) ومن ثم فإن مرحلة التشخيص لم تظهر أي تفوق لأي نموذج من النموذجين (I) ، (II) .

تشخيص النموذج (III)

ذكرنا أن النموذج (III) والذي يمكن صياغته كالتالي :

$$\Delta \log y_t = \delta + (1 - \theta_1 B)\, \varepsilon_t$$

يمثل النموذج الثالث والأخير الذي إخترناه في مرحلة التعريف لتحليل بيانات الحج . وبعد الانتهاء من مرحلة تقدير معالم هذا النموذج ، قمنا بتطبيق نفس الفحوص والاختبارات – السابق إستخدامها لدراسة ملاءمة النموذجين (I) ، (II) – لتشخيص النموذج الذي بين أيدينا . والسطور الأتية تمثل النتائج الرئيسية لهذه الفحوص والاختبارات :

1. لا تتضمن دالة الارتباط الذاتي للبواقي (أو دالة الارتباط الذاتي الجزئي) أي نتوءات .
2. قيمة الإحصاء Q المعدل تساوي 12.5 .
3. الفروق الأولى للبواقي Δe_t لها نموذج MA(1) بمعلمة θ_1 لا تختلف معنوياً عن 1 ومعامل إرتباط ذاتي من الرتبة الأولى لا يختلف معنوياً عن 0.50- .
4. تشتت البواقي حول خط الوسط صفر ثابت تقريباً .
5. رسم البواقي ضد القيم المقدرة لا يظهر أي إرتباط واضح .
6. قيمة إحصاء اندرسون – دارلينج للبواقي تساوي 0.374 وقيمة P المناظرة تساوي 0.389 مما يدل على ملاءمة فرض اعتيادية الأخطاء .
7. إضافة معلمة متوسطات متحركة من الرتبة الأولى ليس له ضرورة .
8. إضافة معلمة انحدار ذاتي من الرتبة الأولى يؤدي الى نتائج سيئة .

وتدل نتائج الفحوص التشخيصية على ملاءمة النموذج (III) لتحليل بيانات الحج أنه ليس هناك فروق جوهرية بين نتائج هذه الفحوص ونتائج الفحوص للنموذجين (I) ، (II) . ومن ثم فإن مرحلة التشخيص لم تظهر أي تفوق لأي نموذج على أي نموذج من النموذجين الأخرين. وعند نهاية مرحلة التشخيص لم نشعر بأي حرج في إختيار أي نموذج من النماذج الثلاثة لتحليل بيانات الحج والتنبؤ بها . ومبدأ الأقتصاد في المعالم يخبرنا بإختيار النموذج الذي

يحتوي على أقل عدد ممكن من المعالم لتمثيل البيانات بشكل ملاءم . وهذا يعني إختبار النموذج (I) أو النموذج (III) . ومن المعروف إحصائياً أن نموذج (I) أبسط في المفهوم والتطبيق من النموذج (III) ، وبالتالي لا نشعر بأي حرج في إختيار النموذج (I) لتحليل بيانات الحج . وبالرغم من ذلك فقد رأينا أنه من الأفضل وضع هذه النماذج تحت التقويم والمفاضلة في مرحلة التنبؤ لمعرفة ما إذا كانت هناك فروق بين قدرات هذه النماذج في إعطاء تنبؤات موثوق بها ومتسقة .

التنبؤ Forecasting

التنبؤ هو المرحلة الأخيرة من مراحل تحليل السلاسل الزمنية ، وهو عادة الهدف النهائي من تحليل السلاسل الزمنية . وتجدر الاشارة هنا الى أننا لا يمكننا الانتقال الى مرحلة التنبؤ إلا بعد الأنتهاء من تطبيق جميع الفحوص والاختبارات الضرورية لتشخيص النموذج (أو النماذج) الذي إختير في مرحلة التعريف والإطمئنان الى أن هذا النموذج قد إجتاز كافة الفحوص والإختبارات بكفاءة عالية . والنموذج الذي إختير بشكل جيد وتم تشخيصه بشكل جيد يؤدي غالباً في النهاية الى تنبؤات موثوق بها ومتسقة من الناحية الإحصائية .

وفي مجال تقويم كل نموذج من النماذج الثلاثة التي تم تشخيصها استخدمت طريقتان هما :

الطريقة الأولى : التنبؤ بالمشاهدات الأخيرة

ولقد استخدمنا هنا أول 20 قيمة من قيم سلسلة الحج (من الفترة ١٣٩٦هـ الى الفترة ١٤١٥هـ) مع كل نموذج من النماذج الثلاثة لإعطاء تنبؤات لأخر خمس سنوات (من الفترة ١٤١٦هـ الى الفترة ١٤٢٠هـ) ثم مقارنة هذه التنبؤات بالقيم الفعلية المتاحة .

الطريقة الثانية : التنبؤ بأول مشاهدة بعد التحديث

وقد استخدمنا هنا أول 20 قيمة من قيم سلسلة الحج (من الفترة ١٣٩٦هـ الى الفترة

١٤١٥هـ) لتقدير معالم كل نموذج ثم التنبؤ بأول مشاهدة (One step ahead forcast) أي التنبؤ بالمشاهدة رقم 21 (الفترة ١٤١٦هـ) ، ثم إستخدم أول 21 مشاهدة لتحديث مقدرات معالم كل نموذج من النماذج الثلاثة واستخدام هذه المقدرات الحديثة في التنبؤ بأول مشاهدة مستقبلية أي التنبؤ بالمشاهدة رقم ٢٢ (الفترة ١٤١٧هـ) ، ثم كررت هذه العملية أربع مرات وقورنت النتائج بالقيم الفعلية .

ولتقويم نتائج التنبؤات المختلفة والمقارنة بين النماذج الثلاثة استخدمنا المعيارين الأتيين :

(1) متوسط الانحرافات المطلقة (MAD)

حيث يعرف هذا المقياس كالتالي :

$$\text{MAD} = \frac{1}{5}\sum_{i=1}^{5}\left|\text{Actual - Forecast}\right|$$

أي متوسط الفروق المطلقة بين التنبؤات والقيم الفعلية .

(2) متوسط نسب الانحرافات المطلقة (MPAD)

ويعرف هذا المقياس كالتالي :

$$\text{MPAD} = \frac{1}{5}\left\{\sum_{i=1}^{5}\left|\frac{\text{ACTUAL - FORECAST}}{\text{ACTUAL}}\right| \text{X } 100\right\}$$

ويعطي جدول (6) القيم المختلفة للمعيارين لكل نموذج مع التنبؤات التي حصلنا عليها باستخدام الطريقة الأولى في التنبؤ . وتعطي الجداول (7) ، (8) ، (9) التنبؤات التي حصلنا عليها باستخدام الطريقة الثانية (التنبؤ بعد التحديث) والقيم المختلفة للمعيارين وذلك للنماذج الثلاثة على الترتيب .

وبفحص نتائج الطريقة الأولى في تقويم التنبؤ أي جدول (6) يمكن أن نستخلص النتائج الأتية :

1. الفروق المطلقة للنموذج (I) أقل بصفة عامة من الفروق المطلقة لكل من النموذجين الأخرين .

2. متوسط الانحرافات المطلقة (MAD) للنموذج (I) يساوي 78940.4 في مقابل 82008.2 للنموذج (II) و 89759 للنموذج (III) . وبالتالي فإن متوسط الانحرافات المطلقة للنموذج (I) يقل عن متوسط الانحرافات المطلقة للنموذج (II) بمقدار 3067.8 ويقل عن متوسط الانحرافات المطلقة للنموذج (III) بمقدار 10818.6 .

3. نسب الانحرافات المطلقة عن القيم الفعليا (P.A.D) للنموذج (I) أقل دائماً من النسب المناظرة لكل من النموذجين الأخرين .

4. متوسط نسب الانحرافات المطلقة عن القيم الفعلية للنموذج (I) يساوي 6.7191 في مقابل 6.9558 للنموذج (II) و 7.5828 للنموذج (III) . وهذا يعني أن النموذج (I) أفضل من النموذج (II) بنسبة 3.5% وأفضل من النموذج (III) بنسبة 12.9% عند استخدام هذا المعيار .

وبفحص نتائج الطريقة الثانية في تقويم التنبؤ أي جداول (7) , (8) , (9) يمكن أن نستخلص النتائج الأتية :

1. متوسط الانحرافات المطلقة (MAD) للنموذج (II) يساوي 78498.6 في مقابل 80969.8 للنموذج (I) و 83975.8 للنموذج (III) . وبالتالي فإن متوسط الانحرافات المطلقة للنموذج (II) يقل عن نظيره في النموذج (I) بمقدار 2471.2 ويقل عن نظيره في النموذج (III) بمقدار 5477.2 .

2. متوسط نسب الانحرافات المطلقة عن القيم الفعلية للنموذج (II) يساوي 6.843 في مقابل 7.06255 للنموذج (I) و 7.31328 للنموذج (III) . وهذا يعني أن النموذج (II) أفضل من النموذج (I) بنسبة 3.2% وأفضل من النموذج (III) بنسبة 6.9% عند استخدام هذا المعيار .

وخلاصة النتائج السابقة أن كل من النموذجين الأول والثاني أفضل من النموذج الثالث ، وأما بالنسبة للنموذجين الأول والثاني فيمكن القول بأنه ليس هناك فروق جوهرية بينهما. وهذا يؤكد إختيارنا الذي سبق أن ذكرناه عند نهاية مرحلة التشخيص وهو أن النموذج (I) يعتبر الأفضل في نمذجة بيانات الحج . ولذلك فقد استخدمنا هذا النموذج في التنبؤ بعدد الحجاج في السنوات الخمس القادمة (من السنوات ١٤٢١هـ الى ١٤٢٥هـ) . ويحتوي جدول (10) على التنبؤات المستقبلية لعدد حجاج وفترات الثقة المناسبة (95% C.I.) لعدد الحجاج في هذه السنوات .

Table(6)

Forecasts of The Last 5 Observations Using The First 20 Observations For The Three Models

Year	Actual	For.(AR1)	For.(AR2)	For.(MA1)	A.D.(AR1)	A.D.(AR2)	A.D.(MA1)	P.A.D.(AR1)	P.A.D.(AR2)	P.A.D.(MA1)
1416	1080465	1051405	1051158	1038947	29060	29307	41518	2.68958273	2.71244325	3.8426048
1417	1168591	1075913	1066684	1056746	92678	101907	111845	7.93074737	8.72050187	9.57092772
1418	1178186	1094870	1090818	1074850	83316	87368	103336	7.07154897	7.41546751	8.77077134
1419	1056730	1116567	1109200	1093264	59837	52470	36534	5.66246818	4.96531754	3.45726912
1420	1267555	1137744	1128566	1111993	129811	138989	155562	10.2410546	10.9651258	12.2726036
Mean					**78940.4**	**82008.2**	**89759**	**6.71908038**	**6.95577119**	**7.58283531**

1)A.D.= ABS. [(Actual - Forecast)]

2) P.A.D.= ABS. [(Actual - Forecast)]*100 /Actual

Table(7)
One Step Ahead Forecast After Updating The Parameters
Of The First Model

Year	δ	φ_1	Actual	Forec.	ABS.DEV.	P.A.D.
1416	0.01146	-0.387	1080465	1051405	29060	2.689583
1417	0.01204	-0.384	1168591	1095637	72954	6.242903
1418	0.01338	-0.38	1178186	1169745	8441	0.71644
1419	0.01351	-0.379	1056730	1211658	154928	14.66108
1420	0.0108	-0.372	1267555	1128089	139466	11.00276
Mean					**80969.8**	**7.062552**

Table(8)
One Step Ahead Forecast After Updating The Parameters
Of The Second Model

Year	δ	φ_1	φ_2	Actual	Forec.	ABS.DEV.	P.A.D.
1416	0.01395	-0.499	-0.289	1080465	1051158	29307	2.712443254
1417	0.01456	-0.496	-0.2889	1168591	1083118	85473	7.314192904
1418	0.01606	-0.487	-0.2751	1178186	1155748	22438	1.904453117
1419	0.01639	-0.482	-0.2717	1056730	1193007	136277	12.89610402
1420	0.01426	-0.487	-0.3016	1267555	1148557	118998	9.387994998
Mean						**78498.6**	**6.843037658**

Table(9)
One Step Ahead Forecast After Updating The Parameters
Of The Third Model

Year	δ	θ_1	Actual	Forec.	ABS.DEV.	P.A.D.
1416	0.00738	0.4856	1080465	1038947	41518	3.842605
1417	0.00793	0.4684	1168591	1082346	86245	7.380255
1418	0.00895	0.4451	1178186	1155898	22288	1.891722
1419	0.0092	0.4381	1056730	1194252	137522	13.01392
1420	0.0078	0.4704	1267555	1135249	132306	10.43789
Mean					**83975.8**	**7.313278**

Table(10)
Five Step Ahead Forecasts Of Pilgrims

Year	Forec.	Lo.Bo.	Up.Bo.
1421	1210484	975116	1502664
1422	1272662	991387	1633741
1423	1284850	953259	1731786
1424	1319537	945938	1840688
1425	1345416	932302	1941586

Appendix

Number Of Pilgrims From Year 1396 To 1420

Year	Pilgrims	Year	Pilgrims	Year	Pilgrims
1396	719040	1406	856718	1416	1080465
1397	739319	1407	960386	1417	1168591
1398	830236	1408	762755	1418	1178186
1399	862520	1409	774560	1419	1056730
1400	812892	1410	827236	1420	1267555
1401	879368	1411	720102		
1402	853555	1412	1012140		
1403	1003911	1413	992813		
1404	969671	1414	995611		
1405	846097	1415	1042374		

المراجع العربية :

- الأمن في الحج ، أ.د. محمود بن محمد سفر ، وزارة الحج .
- الحج منطلق لتحقيق التعاون بين المسلمين ، أ.د. محمود بن محمد سفر ، وزارة الحج .
- منهجية الرؤى المستقبلية للملك عبدالعزيز في ترتيب أوضاع الحج ، أ.د. محمود بن محمد سفر ، وزارة الحج .
- التعليمات المنظمة لشئون الحج ، موسم حج عام ١٤٢٠هـ - وزارة الحج .
- الحج في مائة عام كتاب وثائقي بمناسبة مرور مائة عام على تأسيس المملكة العربية السعودية ، وزارة الحج .

المراجع الأجنبية :

1. Abraham, B. & Ledolter, J. (1983). Statistical Methods for Forecasting, John Willey & Sons, Inc., New York.

2. Anderson, T.W. (1971). The Statistical Analysis of Time Series, John Wiley & Sons, Inc., New York.

3. Bowerman, B.L. & O'connell, R.T. (1987). Time Series Forecasting, 2nd Edition, Duxbury Press, Boston.

4. Box, G.P. & Jenkins, G.M. (1976). Time Series Analysis forecasting and control, Holden-Day, San Francisco.

5. Eatwell, J. Milgate, M. & Newman, P. (1990). Time Series and Statistics: the New Palgrave, Macmillan Press, London.

6. Goldfeld, S.M. & Quandt, R.E (1972). Nonlinear Methods in Econometrics, North Holland Publishing Company, Amesterdam.

7. Granger, C.W.J. (1980). Forecasting in Business and Economics Academic Press, New York.

8. Hamilton & James, D. (1994). Time Series Analysis, Princeton Press, New Jersey.

9. Harry, H.K., & Wallace, E.O. (1981). Introduction to Econometrics: Principles and Applications, 2nd Edition, Harper & Row, New York.